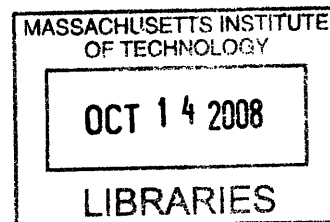


The role of perception in phonotactic constraints:
evidence from Trinidad English.

by

Franz Katzir Cozier

B.S. Neuroscience, Linguistics, and Psychology
University of California, Los Angeles, 2001



SUBMITTED TO THE DEPARTMENT OF LINGUISTICS IN PARTIAL
FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN LINGUISTICS
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

SEPTEMBER 2008

© Franz Katzir Cozier. All rights reserved.

The author hereby grants to MIT permission to reproduce
and to distribute publicly paper and electronic
copies of this thesis document in whole or in part
in any medium now known or hereafter created.

Signature of Author: _____
Department of Linguistics
September 5, 2008

Certified by: _____
Donca Steriade
Professor of Linguistics
Thesis Supervisor

Accepted by: _____
Irene Heim
Professor of Linguistics
Department Head,
Department of Linguistics and Philosophy

ARCHIVES

The role of perception in phonotactic constraints: evidence from Trinidad English.

by

Franz Katzir Cozier

Submitted to the Department of Linguistics on September 5, 2008 in Partial Fulfillment
of the Requirements for the Degree of Master of Science in Linguistics

ABSTRACT

This thesis demonstrates that perception plays a role in computing grammatical well-formedness. This is shown specifically for grammatical constraints on word-final consonant cluster inventories (VC1C2#), with focus on coda cluster simplification in Trinidad English. The first claim of this thesis is that C2 deletion is triggered when VC1C2# is not sufficiently distinguished perceptually from VC1#, by at least one relevant perceptual dimension. The relevant properties that sufficiently distinguish VC1C2# from VC1# are the release burst spectrum, values of F2 transitions to C1 or C2, transitions from C1 into C1, nucleus (Vowel + Sonorant) duration, and high amplitude frication noise. This hypothesis was tested with a perception experiment that measured Standard English speakers' ability to discriminate between attested VC1C2#'s of Standard English and VC1#, with unreleased C2. The result is that those C2's that are significantly less likely to be perceived in the absence of release are the same C2's that are deleted in Trinidad English ($p < .01$). The relevant perceptual dimensions are the ones proposed here: release burst, values of F2 transitions to C1 or C2, transitions from C1 into C2, nucleus duration, and high amplitude frication noise.

The second claim of this thesis is that speakers encode this perceptually based difference between simplified and preserved clusters in their grammars. Namely, speakers neutralize VC1C2# and VC1# to VC1# where there is subminimal perceptual contrast between VC1C2# and VC1#. In order to test this second hypothesis, the perceptual discriminability of VC1C2# and VC1#, for unattested clusters of English, was established in a perception experiment. Some unattested VC1C2's were significantly more perceptually distinct from VC1 than others ($p < .01$). This was predictable based on

the relevant properties: release burst spectrum, values of F2 transitions to C1 or C2, transitions from C1 into C2, nucleus duration, and high amplitude frication noise. In an affix stripping experiment, it was then discovered that TE speakers do not simplify unattested clusters (like mg# or mk#) across the board, but rather they simplify all clusters as a function of the perceptual difficulty involved in discriminating VC1C2# from VC1#.

Table of Contents:

Introduction	p. 5
1.0 Cluster simplification in TE	p. 7
2.0 Hypothesis I: C2 deletion where C2 is not cued.	p. 10
3.0 Testing HI with a perception experiment.	p. 17
4.0 Hypothesis II: Encoding of perceptual dimensions relevant to preservation or simplification of C2.	p. 22
5.0 Testing HII with a production experiment.	p. 26
6.0 OT Analysis: Modeling perceptual factors into the grammar.	p. 38
7.0 Word-final coda cluster simplification crosslinguistically.	p. 46
8.0 Conclusion/Discussion.	p. 50

INTRODUCTION*

The relative salience of speech sounds is an important factor in explaining phonological behavior (Kohler 1990; Hura, Lindblom and Diehl 1992, Ohala 1990, and many others). An important factor in understanding this relationship between perception and phonology is the effect of different acoustic environments on the perceptibility of speech sounds (cf. Steriade 1997).¹ Steriade (2001a, b) proposes that there is a component of the grammar called the P-map, which is a set of statements about absolute and relative perceptibility of different contrasts across the different contexts where they might occur. For example, the P-map may contain a speaker's knowledge that a voicing contrast between two stops in word-final position is better perceived after a vowel (VC1# vs. VC2#) than after a consonant (CC1# vs. CC2#).

Steriade also claims that the perceptual salience of a segment is the degree to which that segment is not confusable with zero. Perceptual salience of a speech sound is a function of the quantity and quality of auditory cues that signal the presence of that sound. For example, Wright (2004) proposes that several sources of cues for consonants include second formant transitions in neighboring vowels, stop release bursts, and frication noise.

These and other work have established that perception does play a role in phonology. However, there is still a need for more evidence to determine whether this relationship is a historical one or a synchronic one. It could be that less perceptible sounds are less stable because they are more likely to be misheard or not heard at all (cf. Blevins 2004, Hyman 2001, Ohala 1990, Barnes 2002, Boersma 1998). This paper demonstrates that perception plays a role in computing grammatical well-formedness. This will be shown specifically for grammatical constraints on word-final consonant cluster inventories (VC₁C₂#).

* I am thankful to Donca Steriade for guidance and support in this endeavor. I'm also thankful to audiences at the Manchester Phonology Meeting 2007 for the opportunity to present some of the results of this paper. Thanks also to WCCFL 2008, and the Consonant Clusters and Structural Complexity Workshop, Munich 2008 for the opportunity to presents these results. Thanks also to Adam Albright and Edward Flemming for comments on earlier versions of this paper.

¹ Steriade (1997) shows that voicing contrasts are more common in contexts where more cues to voicing are available.

The set of word final clusters that simplify ($VC_1C_2\# \rightarrow VC_1\#$) is remarkably similar across different languages. Languages that simplify final codas include Quebec French (QF), Trinidad dialectal English (TE), African American English (AAE), Cameroon English, Tejano English, Jamaican English, Catalan, and Korean (cf. Côté 2004, 2000; Green 1992; Bobda 1994; Bayley 1995; Akers 1981, Meade 2001, Patrick 1991; Mascaro 1978; Cho 1999). Most languages that ban word-final clusters also lack release of the final consonant (cf. Archambault & Dumochel 1993: QF; Katzir Cozier 2008: TE; Wheeler 1979, Mascaro p.c.: Catalan; Martin 1951, Huh 1965, McCawley 1967, Kim 1971, Kim-Renaud 1974, Chung 1986, Sohn 1987, Baek 1991, Lee 1994, Kim 1998, Rhee 1998, among others: Korean). Secondly, it is always the less perceptible consonant (C2), and not C1 that deletes. The best cue to a consonant apart from its release is found in the formant transitions on neighboring vowels; C1 is adjacent to the vowel so its features are, generally speaking, better cued than those of C2.

The lack of C2 release in many of the languages that simplify word-final clusters, and the fact that such simplification is observed for the less perceptible of the two final consonants (C2), make it seem likely that cluster simplification is related to the decreased perceptibility of C2.

This paper makes two central claims. The first is that deletion of C2 is triggered when the distinctiveness of the contrast between $VC_1C_2\#$ and $VC_1\#$, as a function of the phonetic cues differentiating them in a word-final context, falls below a particular threshold. The second claim is that speakers encode this perceptually based difference between simplified and preserved clusters in their grammars.

This paper is divided into eight sections. In the first section, cluster simplification data is presented for Trinidad English. In the second section, the hypothesis is presented that deletion of C2 is triggered by the subminimal contrast between $VC_1C_2\#$ and $VC_1\#$, due to insufficient cues to C2. In the third section, this hypothesis is tested with a perception experiment that measures Standard English speakers' ability to discriminate between attested $VC_1C_2\#$'s of SE and $VC_1\#$, with unreleased C2. The result is that those C2's that are significantly less likely to be perceived in the absence of release are the same C2's that are deleted in TE. In the fourth section, the perceptibility of C2 for unattested clusters is established in a perception experiment. This is to test my second hypothesis,

namely that speakers encode the perceptually based difference between simplified and preserved clusters, and neutralize VC₁C₂ and VC₁ to VC₁ where there is subminimal contrast between VC₁C₂# and VC₁#. In section five, this second hypothesis is tested with an affix stripping experiment. The experiment compared final clusters (C₁C₂) that are sufficiently distinct from C₁# (e.g. pʃ-p) with clusters that are more confusable (pf-p).² The result is that TE speakers do not simplify unattested clusters (like mg# or mk#) across the board, but rather they simplify all clusters as a function of the perceptual difficulty involved in discriminating a cluster from its simplified form. The sixth section presents a model of how perception can be integrated into the computation of grammatical well formedness.

The focus of this paper is going to be Trinidad English, but the results should be widely applicable to any language without final release. The crosslinguistic implications of the current results are presented in section seven.

1.0 CLUSTER SIMPLIFICATION PATTERNS IN TRINIDAD ENGLISH

The inventory of allowable word-final clusters is smaller in Trinidad English than in Standard English (SE). Simplification is characterized by loss of the second consonant in a cluster. It is always C₂ that deletes (VC₁C₂# → VC₁#). Medial clusters don't delete in TE.

- (1) mandate [mændet], *[mænet]; banter [bæntʌ], *[bænʌ]; mister [mɪstʌ], *[mɪsʌ];
whisper [wɪspʌ], *[wɪsʌ]; actor [aktʌ], *[akʌ]; mixer [miksʌ], *[mikʌ].

Some of the clusters shown in (1), namely [nd], [st], and [kt] simplify word finally even though they do not simplify medially. The other clusters, namely [nt], [sp], and [ks] are preserved both medially and word finally. The data on the pattern of cluster simplification in word final position will now be presented. The data in this section

² Strident fricatives are more perceptually salient than non-strident fricatives (cf. Chomsky & Halle 1968, Miller & Nicely 1955).

focuses on TE. Simplification is not optional in TE. Word-final stops are unreleased in TE.

1.1 Factors in cluster simplification

In TE voicing differentiates the homorganic nasal-stop and homorganic liquid-stop coda clusters that are simplified from those that do not. Homorganic sequences of sonorant-voiced stop clusters are simplified while sonorant-voiceless stop clusters are preserved. This is shown in (1) for both nasal-stop and liquid-stop clusters.

(1) Patterns of C-deletion in TE: Sonorant-stop coda clusters

Son = Sonorant; T = voiceless stop

<...> = deleted Consonant; C' = unreleased stop

	(a) VSon ₁ T ₂ #	(b) VSon ₁ d ₂ #
SE	[vɛnt] “vent” [kolt] “colt”	[wind] “wind” [tʃajld] “child”
TE	[vɛntʰ] “vent” [koltʰ] “colt”	[wɪn] <d> “wind” [tʃajl] <d> “child”

This pattern, where the final voiced stop is deleted in homorganic clusters and the voiceless stop preserved, as well as the upcoming simplification patterns in (3), are ubiquitous crosslinguistically. We discuss this aspect in section 7.

For equally voiced C1 and C2's in liquid-stop clusters, heterorganicity makes a difference. The coda cluster [lb], a liquid followed by a heterorganic stop, is preserved, even though /ld/ is simplified (cf. 1):

(2) C2-preservation in TE: heterorganic liquid-stop coda clusters

	Vl ₁ b ₂ #
SE	[bʌlb] “bulb”
TE	[bʌlbʰ] “bulb”

This pattern will be shown later to be true of other languages, including QF where homorganic /ld/ is simplified and heterorganic /lb/ and /lg/ are preserved. Note that [lg] is

not present in Standard English; it is missing in TE for this reason. It should also be noted that /r/ does not exist at all in codas of TE. TE is also a non-rhotic dialect of English.

In (1), it was shown that voiceless stops are preserved after all sonorants. In (3) it is shown that C2 is lost after all obstruents.

(3) Patterns of C-deletion in TE: Obstruent-obstruent coda clusters

	(a) VT ₁ T ₂ #	(b) VF ₁ t ₂ #
SE	[akt] “act” [æpt] “apt”	[lɪft] “lift” [rost] “roast” [tæsk] “task”
TE	[akʰ] “act” [æpt] “apt”	[lɪf] “lift” [ros] “roast” [tæs] “task”

However, /sp/ is preserved in TE, as shown in (4). This is in contrast with /st/ and /sk/, which are simplified (cf. (3b)). Labiality makes a difference here.

(4) C2-preservation in TE: /sp/

	Vs ₁ p ₂ #
SE	[kɾɪsp] “crisp” [lɪsp] “lisp”
TE	[kɾɪspʰ] “crisp” [lɪspʰ] “lisp”

The presence or absence of frication noise in C2 also has an effect on coda cluster simplification. All fricative and affricate C2's are always preserved, after sonorants and obstruents, regardless of the voicing of the fricative or affricate.

(5) C2-preservation in TE: Sonorant-fricative and Obstruent-fricative coda clusters

	(a) VSon ₁ F ₂ #	(b) VT ₁ s ₂ #
SE	[wʌns] “once”, [bɪndʒ] “binge”, [lɛnz] “lens”, [bʌldʒ] “bulge”, [gɔlf] “golf”	[tʃreps] “traipse” [bɒks] “box”
TE	[wʌns] “once”, [bɪndʒ] “binge”, [lɛnz] “lens”, [bʌldʒ] “bulge”, [gʌlf] “golf”	[tʃreps] “traipse” [bʌks] “box”

The following factors play a role in the preservation of a cluster: voicing, heterorganicity, labiality, and frication noise.

One could posit several rules to treat these phenomena. For instance:

(6) (a) [+coronal, +voi] → Ø/ [+coronal, +voi] ____#

(b) [-voi, -cont, -lab] → Ø/ [-voi] ____#

However, this kind of approach misses the generalization that deletion of C2 is predictable from the fact that C2 is unreleased, in TE and other languages with final unrelease. The analysis in (6) does not explain why C2 is deleted. The approach laid out in the next section captures the fact that deletion is triggered by one underlying factor, namely the perceptual difficulty in distinguishing a consonant cluster from its simplified counterpart in the absence of C2 release.

2.0 HYPOTHESIS I: C2 DELETION WHERE C2 IS NOT CUED

I hypothesize that a word-final cluster surfaces in any language only if it is sufficiently perceptually distinct from its simplified counterpart.³ Accordingly, we will compare throughout VC1C2# with VC1#. Sufficient perceptual distinctness is expressed as a difference in any one of the following properties (cf. Wright 2004).

(7) Relevant properties that sufficiently distinguish VC1C2# from VC1#

- (a) High amplitude frication noise: “tax” **ta[ks]** – “tack” **ta[k]** vs. “tact” **ta[kt]** – “tack” **ta[k]**
- (b) F2 transitions in C1: “bulb” **bu[lb]** – “bull” **bu[l]** vs. “bold” **bo[ld]** – “bowl” **bol]**
- (c) Transitions from /s/ to a labial (cf. Munson 2001): “grasp” **gra[sp]** – “grass” **gra[s]** vs. “best” **be[st]** – “Bess” **[bes]**
- (d) Nucleus duration: “ant” **a[nt]** – “an” **a[n]** vs. “and” **a[nd]** – “an” **a[n]**; “bolt” **bo[lɪt]** – “bowl” **bo[l]** vs. “bold” **bo[ld]** – “bowl” **bo[l]**
- (e) Release burst: “pact” **pa[kt]** – “pack” **pa[k]**

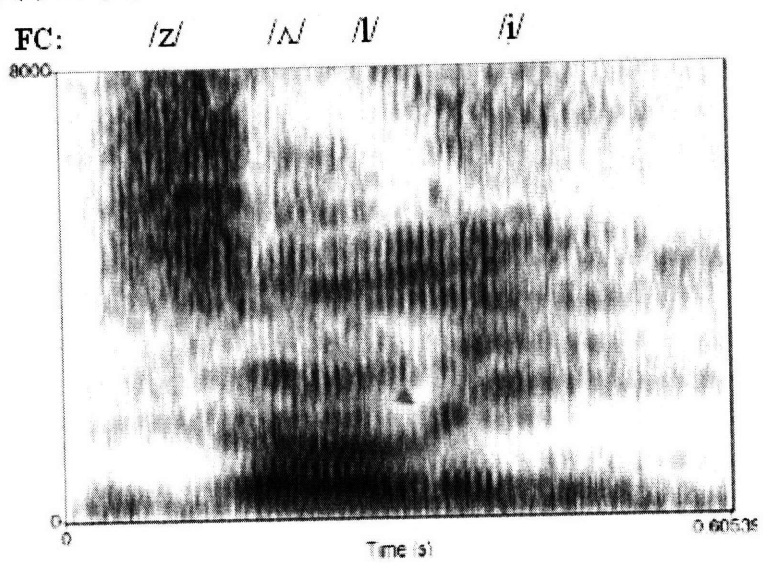
The first relevant property that sufficiently distinguishes VC1C2# from VC1# is high amplitude frication noise, which insures sufficient perceptual distinctness between pairs like *tax* and *tack* in English but not between pairs like *tact* and *tack*.

Example (7b) illustrates the fact that the F2 transitions in C1 sufficiently distinguish VC1C2# from VC1# in pairs like *bulb* and *bull*. This cue is missing in pairs like *bold* and *bowl*. This is shown in (8) where F2 values in /l/ are identical between the final cluster /ld/# and final /l/#, but F2 values in /lb/# are different from the ones in /l/#.

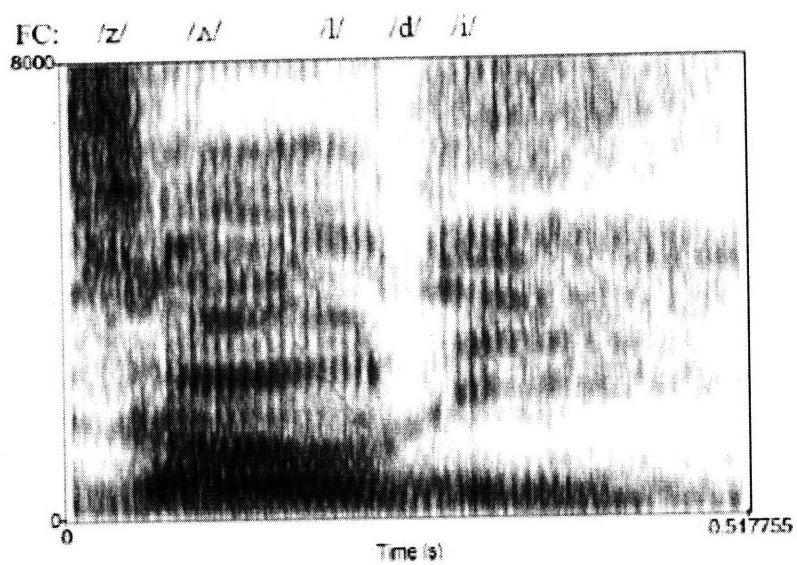
The spectrograms in (8) are nonce words of TE recorded in isolation. The purpose of the recording was to see how distinguishable VC1C2# is from VC1#, given the cues listed in (7). However I propose that simplification occurs for some clusters in word-final position because these C2’s are not cued in that position. In order to establish acoustically that C1C2’s that simplify are not sufficiently distinct from C1, and that preserved C1C2’s are sufficiently distinct from C1, comparisons need to be made in a context where all clusters can occur. To that end, clusters were compared acoustically in a medial context, where all clusters are preserved, as predicted by Hypothesis I above and (7), due to the fact that C2 is cued by the inevitable release into a following vowel. I am making the assumption that if we ignore the release cues for C2, then the acoustic differences established between comparisons of different VC1C2V#’s and VC1V#, would also hold for VC1C2# and VC1#, given that only C2 is varied across all comparisons.

³ I should note that the generalizations I have made here are about clusters that abide by the sonority sequencing principle, i.e. where there is no serious rise in sonority in the final cluster. Sonority increasing clusters may be subject to simplification even if VC1C2# is sufficiently perceptually distinct from VC1#

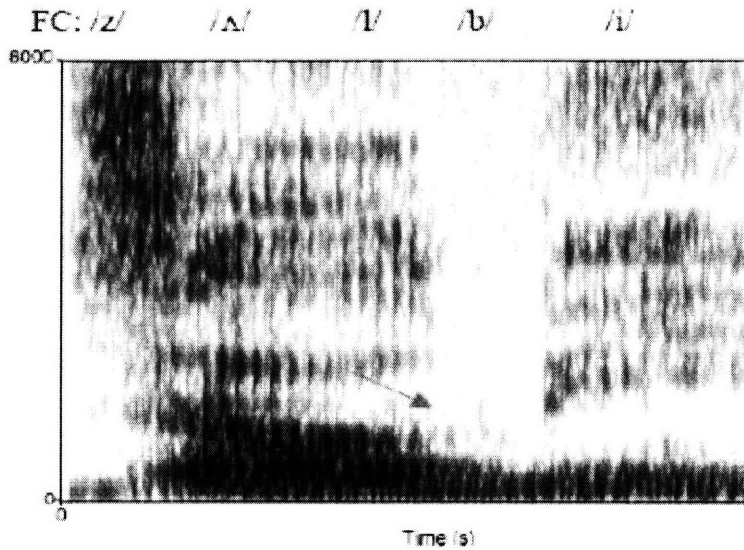
(8) (a) *zully* (recorded in isolation)



(b) *zuldy* (recorded in isolation)



(c) *zulby* (recorded in isolation)

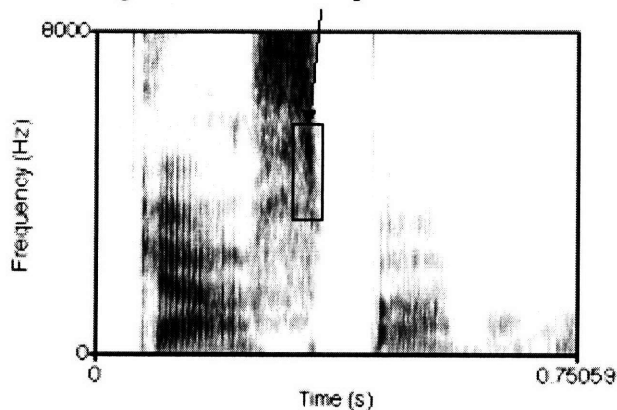


In (8), we see that when we ignore the release cues for C2, then medial LD (b) is less distinct from L (a) than medial LB (c) is from L. This is because heterorganic LB has a falling F2 in C1.

The third property that insures the perceptual contrast between VC1C2 and VC1 are the transitions from /s/ into a labial. This distinguishes pairs like *grasp* and *grass* but not pairs like *best* and *Bess*. For /sp/, it has been observed that there is a broad peak in the /s/ spectrum, centering below 8000 Hz, that drops rapidly toward the end of the /s/ when it is followed by labial /p/, as shown in (9) for TE (cf. Munson 2001).

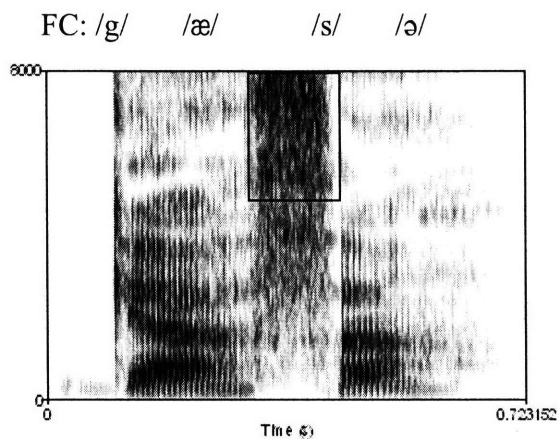
(9) *gasper* (TE); “say ___ please”

FC: /g/ /æ/ /s/ /p/ /a/

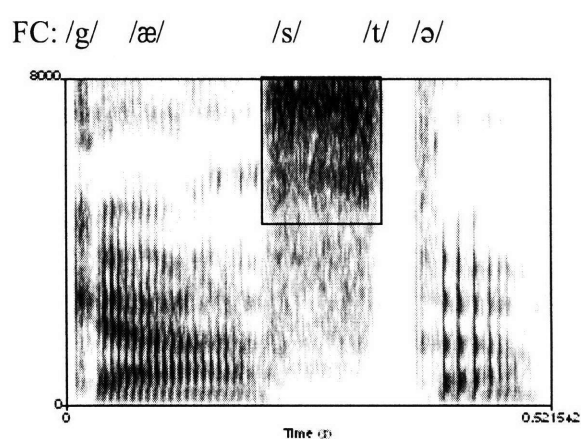


This transition cue to a labial C2 in C1 is not observed for /s/ or /st/ as shown in (10a) and (10b) respectively.

(10) (a) *gasser* (TE); “say ____ please”

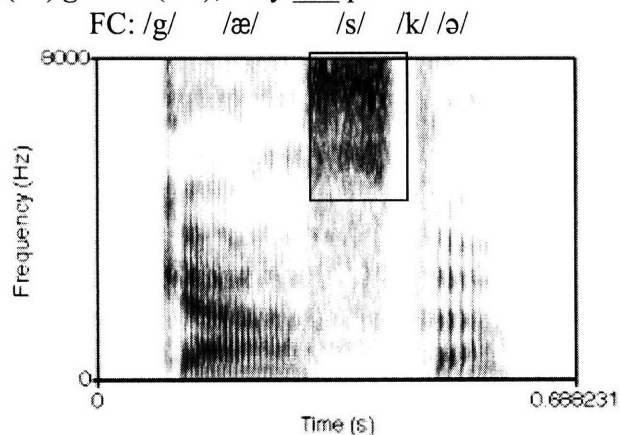


(b) *gaster* (TE); “say ____ please”



Nor, is this transition cue observed for /sk/ as shown in (12).

(12) *gasker* (TE); “say ____ please”



The fourth property that insures perceptual distinctness of VC1C2-VC1 is nucleus duration. The nucleus duration distinguishes minimal pairs like *ant* and *an* but not minimal pairs like *and* and *an*. Likewise, the nucleus duration distinguishes pairs like *bol* and *bow* but is identical in pairs like *bold* and *bow*. It is known that in some varieties of English, vowel shortening occurs preceding a voiceless stop but not a voiced one (cf. Naeser 1970).

To verify this for TE, nonce tokens were recorded in isolation from 2 native speakers of TE. For each speaker, each token was recorded 5 times. The average duration was computed. Since simplification does not occur intervocalically, duration effects of C2 for all clusters could be compared in this context independently of whether they simplify or not. The assumption is being made that the duration effects or lack thereof, as observed in an intervocalic context, would also be true word finally. For tokens where C1 is a sonorant, duration measurements are of the preceding vowel and this sonorant. For tokens where C1 is a stop, duration measurements are of the preceding vowel.

(13) Average nucleus duration in TE as a function of voicing of C2.

VC ₁ C ₂ V- VC ₁ V	Δ (in ms)	C2 status in TE
<u>VND</u> V-V <u>N</u> V	11	Reduced
<u>VNT</u> V-V <u>N</u> V	44	Preserved
<u>VLD</u> V-V <u>L</u> V	3	Reduced
<u>VL</u> T <u>V</u> -V <u>L</u> V	57	Preserved
<u>VTT</u> -VT	10	Reduced

Stimulus list: VND: pender; VNT: penter, penker; VN: penner; VLD: pelder; VLT: pelter, pelker; VN: peller; VTT: pefter⁴, peckter, pepter; VT: peffer, pecker, pepper.

The degree of shortening for clusters with voiceless C2 (NT and LT) is significantly greater than for clusters with voiced C2 (ND and LD). Furthermore, there is no cumulative shortening. A vowel followed by two voiceless obstruents is as short as one followed by a single voiceless obstruent. Therefore the duration of the vowel does not distinguish Vkt from Vk or Vst from Vs.

It should be noted here that while /lt/ is perceptually distinct from /l/, due to nucleus duration, heterorganic /lk/ is perceptually distinct from /l/ because of both nucleus duration, like /lt/, and transitions in C1 like /lg/and /lb/. The spectrum of the release burst distinguishes SE pairs like *pact* and *pack*, where C2 is released. However, this property is missing in simplifying languages like Trinidad English, Quebec French, and Catalan.

⁴ Measurements for /ft/# were consistent with a previous study showing that there is not V shortening before /ft/# clusters even though there is shortening before a voiceless stop in English (cf. Naeser 1970).

Hypothesis I states that a VC1C2# surfaces only if it is sufficiently perceptually distinct from VC1#, due to relevant properties: Release burst spectrum, values of F2 transitions to C2, transitions from /s/ to labial, nucleus (V+R) duration, and high amplitude frication noise. This predicts that in the absence of release as a cue, clusters that are preserved in languages like Trinidad English are sufficiently perceptually distinct from their respective simplified forms, due to other relevant non-release cues. This is illustrated in (14).

(14) Relevant properties perceptually distinguishing VC1C2# from VC1#

Cluster Contrast □ = Simplified	V/Son Dur	Trans In L	Trans In s	Noise	Burst (SE)
VL ₁ T ₂ # - VL ₁ #; VN ₁ T ₂ # - VN ₁ #	+				+
VL ₁ b ₂ # - VL ₁ #		+			+
VL ₁ d ₂ # - VL ₁ #; VN ₁ d ₂ # - VN ₁ #					+
VS ₁ p ₂ # - VS ₁ #			+		+
VS ₁ t ₂ # - VS ₁ #; VS ₁ k ₂ # - VS ₁ #; VT ₁ T ₂ # - VT ₁ #					+
T ₁ F ₂ # - VT ₁ #				+	

Even when the release is missing, some VC1C2's sequences provide a sufficient contrast with VC1, due to nucleus duration, transitions in L, transitions in /s/, and frication noise. It is precisely those C2's that are not cued in the absence of release, which are deleted in TE. Clusters that are simplified in TE are shaded in (14). As indicated previously, sonorant-voiceless stop clusters are cued by nucleus duration, [lb] clusters are cued by transitions in /l/, [sp] clusters are cued by transitions in /s/, and fricative C2's are cued by frication noise. These clusters are preserved in TE. However, in the absence of release the C2's of [ld], [nd], [st], and [sk] clusters have no relevant cues. These clusters are simplified in TE.

The table in (14) shows that for a language like TE, where the release burst is not an available property providing perceptual contrast between VC1C2 and VC1, there will always be a subset of VC1C2's that have no other relevant properties that sufficiently contrast them with VC1. This subset includes [ld], [nd], [st], [sk], and [TT]. I am unaware of any languages without final release where this subset of clusters is preserved. This is true for non-C2-releasing Trinidadian English, African American English, Quebecois French, Korean, and Catalan (cf. Côté 2004, 2000, Green 1992, Mascaro 1978).

However, for languages that release C2, there is a greater tendency to preserve this subset of clusters. For example, this is true of Standard French and SE. Even though SE releases final consonants variably, there is a greater tendency to release C2 in VC1C2# (cf. Kang 2003, Albright 2006). We find that SE does preserve [ld], [nd], [st], [sk], and [TT] clusters.

3.0 TESTING HYPOTHESIS I WITH A PERCEPTION EXPERIMENT

In order to test the hypothesis that deletion of C₂ is triggered when the distinctiveness of VC₁C₂# and VC₁# falls below a particular threshold, a perception experiment was done. The task was to discriminate a cluster with dereleased C2 from its simplified counterpart. If Hypothesis I (section 2.0) is correct, then VC1C2#'s should be perceptually indistinct from VC1#, if there are no relevant properties left, contrasting them with VC1# in the absence of release. The other relevant properties are values of F2 transitions to C1 or C2, transitions from /s/ to a labial, Nucleus (V+R) duration, and high amplitude frication noise. By editing out releases from C2's of the clusters represented in (14), and measuring how well people discriminate these clusters from VC1#, Hypothesis I (section 2.0) can be tested.

3.1 Participants

5 male and 4 female native American English speakers at MIT volunteered for the experiment. None of the participants reported any hearing problems.

3.2 Stimuli

Nonce monosyllabic VC1C2# and VC1# stimuli were recorded from a male speaker of Standard American English (SAE), representing each word-final coda cluster type of English (cf. 15). Refer to the list of these stimuli in Appendix (experiment 1). All these syllables were produced in isolation and recorded at 44,100 Hz using a Shure SM10A head-mounted microphone and a high quality digital recorder, in the sound attenuating

booth of the phonetics lab of the MIT Linguistics Department. Each nonce word was recorded three times. During the recording, the speaker was asked to pause after each word, so that the words would not be read like a list but would have comparable intonation throughout. Out of the three repetitions of each nonsense word, a single token was chosen so as to make up a stimulus set where the intonation shape would be uniform throughout. Selected words were also inspected to determine that C2 was released. There were audible releases for C2 in all the stimuli. In order to have uniform loudness, tokens were equalized using PRAAT script *rms equalize* (Beckers, Gabriel: www.bio.leidenuniv.nl/~eew/G6/staff/beckers/beckers.html). These syllables were then subject to waveform editing in PRAAT, where the final releases were removed so that word final C2's were *dereleased*, creating stimuli comparable to the unreleased VC1C2# strings that speakers of TE would have been exposed to prior to the changes that resulted in simplification of coda clusters.

For the experiment, there were 22 sets of nonce syllables, each representing a different cluster type of English and the simplified counterparts to all these cluster types:

- (15) a. CVnt', CVnd', CVn, CVmp', CVm
- b. CVlt', CVld', CVlk', CVlb', CVl, CVlm, lv
- c. CVsp', CVsk', CVst', CVs
- d. CVkt', CVks, CVk, CVfs⁵, CVft, CVf

For each set there were 5 variants of the CV portion of the syllable. These were kept the same across all sets (cf. Appendix 1, experiment 1). Each token had a dereleased C2. There were 16 cluster types, hence 80 target stimuli and 6 target simplified forms. Additionally there were 460 fillers (cf. Appendix, experiment 1).

3.3 Procedure

The task was binary forced-choice identification. The dereleased VC1C2'# stimuli and their VC1# counterparts were presented aurally to SE subjects. In (15a), the contribution of the shortening effect is being tested. In (15b) the contribution of the shortening effect

⁵ This cluster only appears when a word ending with /f/ is marked for the plural for both TE and QF.

(lt-l vs. ld-l) and transitions in /l/ (lb-l, lv-l⁶, lm-l vs. ld-l) is being tested. In (c) the effect of the transition in /s/ is being tested. In (d) the contribution of high amplitude frication noise is being tested. There were two presentations of each set. Testing was computer facilitated with the use of the software program PsyScope. Stimuli were presented on Sony MDR-7506 headphones at a comfortable listening level. Presentation of all stimuli was randomized by Psyscope for each subject. Participants were free to take as much time as they wanted in producing a response.

After hearing each stimulus, participants were given two choices on a screen (VC1C2# or VC1#). One choice appeared on the left and the other on the right. The order of the choice of VC1C2# or VC1# on the screen was different for each of the two randomized presentations of each token. Two buttons on a keyboard were each marked with a different color and each corresponded, by relative position, to one of the two choices on the screen. Participants were instructed to press the button that corresponded to the choice that matched best with what they had just heard. Tapping one of these keys triggered presentation of the next stimulus in the PsyScope-generated randomization. There were 16 cluster types, hence 160 presentations of target stimuli and 160 presentations of analogous singletons (cf. Appendix, experiment 1).

The perceptual difference between VC1C2# and VC1# for each cluster type was quantified using a sensitivity measure (d'), which takes into account both hits and false alarms (Macmillan and Creelman 2005). Hits were defined as the rate at which VC1C2^{*}#, with dereleased C2, was identified as VC1C2#. False alarms were defined as the rate whereby VC1# was identified as VC1C2#. All significance values reported are from two-tailed, paired T tests.

3.4 Results

Some C1C2 sequences are significantly better discriminated from C1, as indicated by significantly higher d' values. These coincide with the coda clusters that are preserved in TE. The table in (16) displays the data. The relevant factors for preservation are indeed

⁶ There is also a contribution from frication noise here though not with not as high an amplitude as for strident fricatives.

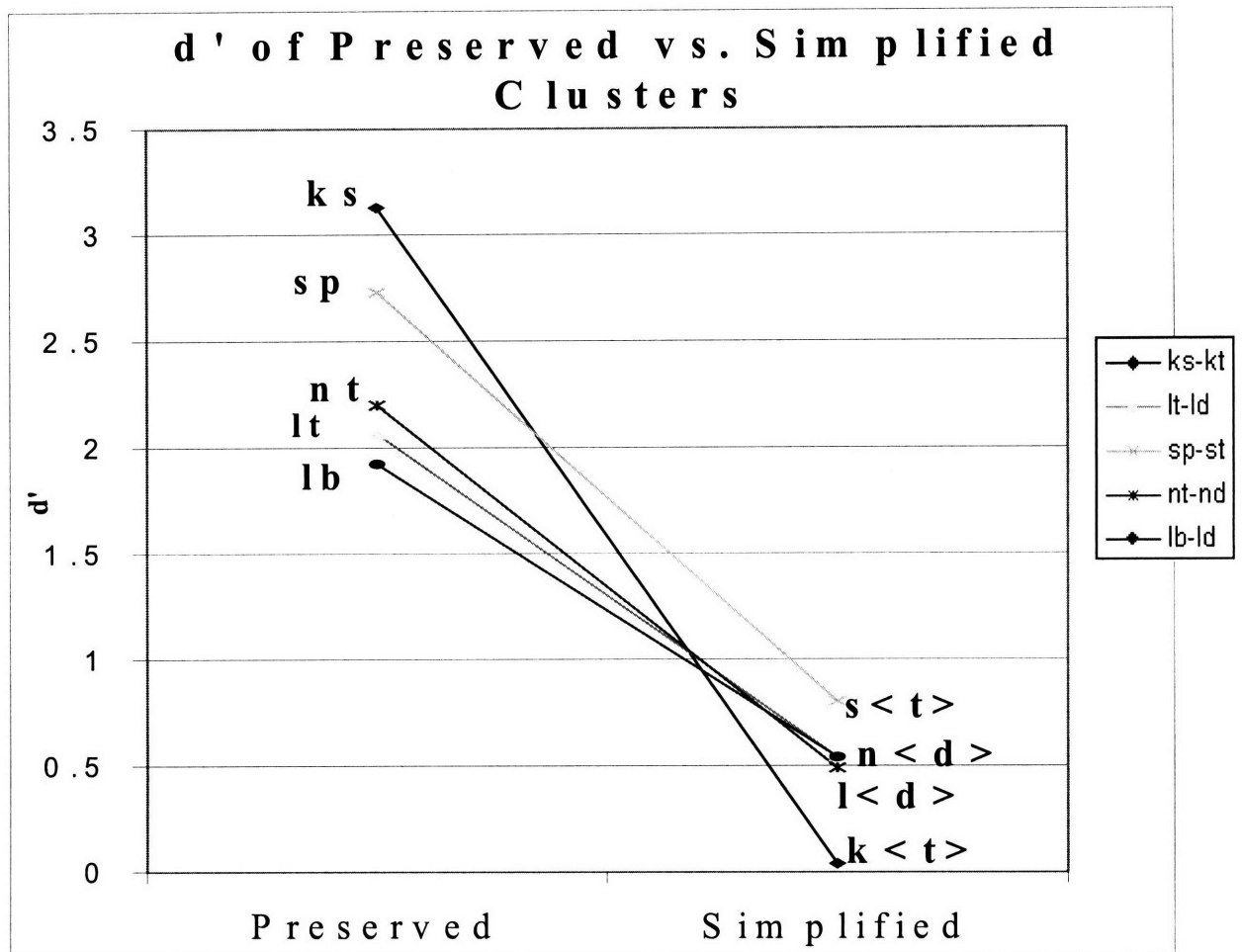
frication noise in C2 (as in /ks/#), transitions to C2 in C1 (as in /lb/# and /sp/#), and nucleus shortening (as for clusters NT# and LT#).

(16) d' values for distinguishing attested VC1C2#'s of English from VC1#.

VC1C2#-VC1 contrast	d'	VC1C2# status in TE
[kt] – [k]	0.040367	Simplified
[ks] – [k]	3.128238	Preserved
[ft] – [f]	0.955046	Simplified
[fs] – [f]	3.003744	Preserved
[st] – [s]	0.802394	Simplified
[sk] – [s]	1.272233	Simplified
[sp] – [s]	2.727469	Preserved
[ld] – [l]	0.541187	Simplified
[lt] – [l]	2.064085	Preserved
[lk] – [l]	2.19882	Preserved
[lb] – [l]	1.923658	Preserved
[lm] – [l]	3.200458	Preserved
[lv] – [l]	2.998623	Preserved
[nd] – [n]	0.490018	Simplified
[nt] – [n]	2.19882	Preserved
[mp] – [m]	1.901492	Preserved

Clusters that are preserved in TE (/ks/#, /sp/#, /nt/#, /mp/#, /lt/#, and /lb/#, /lv/#, /lm/#, /lk/#) have d' values for contrast with their respective C1#'s, that are significantly higher than that of minimally different clusters (/kt/#, /st/#, /nd/#, and /ld/#), which are simplified ($p < .01$). The d' value for distinguishing the pairs {lv, lm, lk, lb, lt}-l is significantly greater than d' for distinguishing ld-l. The d' for distinguishing /fs/# from /f/# is significantly greater than that for /ft/#-/f/. The d' for distinguishing /sp/# from /s/# is significantly greater than that of the pairs {sk, st}-s. These comparisons are illustrated in (17)

(17) d' Significantly Greater for Preserved Clusters than Simplified Clusters



Hypothesis I (section 2.0) is borne out by these data. Deletion of C2 is triggered when the distinctiveness of VC1C2# and VC1# falls below a particular threshold. The only relevant perceptual dimensions for cluster preservation are the ones that I have asserted here:

(18)

(a) Release burst spectrum

(b) Values of F2 transitions to C2 (for [lb], [lm], [lv], [lk])

(c) Transitions from /s/ to labial (for [sp])

(d) Nucleus (V+R) duration⁷ (for [nt], [mp], [lt], [lk])

(e) Hi amp frication noise (for [ks], [fs])

4.0 HYPOTHESIS II: ENCODING OF PERCEPTUAL DIMENSIONS RELEVANT TO PRESERVATION OR SIMPLIFICATION OF C2

Now let's go back to the question of what it is that people know about cluster simplification: Does this knowledge consist of a memorized list of clusters or are the perceptual factors responsible for simplification encoded in the grammar? Does this perceptual analysis extend to clusters that have not been observed in SE?

I claim that speakers do encode in their grammars the perceptually based difference between simplified and preserved clusters, and that this difference extends to clusters that are not found in either TE or SE. This claim will be tested by having TE speakers judge the grammaticality of unattested clusters.

The hypothesis that VC1C2# is grammatically well formed, if it is sufficiently perceptually distinct from VC1#, can be applied to clusters unattested in English (cf. 19). The following clusters have been investigated: [mk], [mg], [lg], [zb], [zd], [kp], [md₃], and [pʃ].

(19) Predictions for simplification of unattested clusters in languages with unreleased C2.

Cluster Contrast □ = Simplified	V/Son Dur	Trans. In L	Trans. In s	Noise	Burst (SE)
Vm ₁ k ₂ # - Vm ₁ #	+				+
Vm ₁ g ₂ # - Vm ₁ #					+
Vl ₁ g ₂ # - Vl ₁ #		+			+
Vz ₁ b ₂ # - Vz ₁ #			+		+
Vz ₁ d ₂ # - Vz ₁ #					+
Vk ₁ p ₂ # - Vk ₁ #					
Vm ₁ d ₃ ₂ # - m ₁ #				+	
Vp ₁ ʃ ₂ # - Vp ₁ #					

⁷ These results here for NT versus ND clusters are corroborated by results from Kaplan (2006) showing that [nd] is harder to perceive word-finally (confused with [n]) than word-internally (intervocally). Cluster NT was shown to be confusable with [n] to a lesser extent word-finally. In the experiment, white noise was added after recording of stimuli. A post-hoc analysis revealed that tokens of VNT had significantly shorter vowels than VND or VN.

Cluster [mk] should be distinct from /m/ due to nucleus shortening before a voiceless stop. Cluster [lg], being heterorganic, should be distinct from /l/ due to values of F2 transitions in C1, as described previously for heterorganic [lb].⁸ It was verified for /lg/#, as for /lb/#, that F2 values in /l/ is different between /lg/# and final /l/#, even though F2 values in /l/ are identical between /ld/# and final /l/#. Cluster [zb] should be distinct from /z/ due to transitions from /z/ to the labial, as seen from /s/ to the labial. This labial transition, demonstrated for [sp], was also verified acoustically for [zb]. For clusters, [md₃] and [pʃ], C1C2 is perceptually distinct from C1 because of the high amplitude frication noise of C2.

However, in the absence of the release burst as a cue, clusters [mg], [zd], and [kp] are not predicted to have any relevant properties making VC1C2 perceptually distinct from VC1, as shown in (19). [Vmg] sequences do not undergo shortening because C2 is not voiceless. Nor does [mg] have any relevant properties making it perceptually distinct from VC1, as we saw for [lg]; only transitions in a vowel or liquid matters. Like [st], cluster [zd] does not have a transition into a labial C2 like [sp]. For cluster [kp], no cumulative vowel shortening is expected, distinguishing [kp] from /k/, as was seen for [kt] in TE. Nor is C2 cued by frication noise. If hypothesis I (section 2.0) can be extended to unattested clusters, due to the encoding of perceptual differences between simplified and preserved clusters in TE, then we predict that clusters [mg], [zd], and [kp] would be more likely to be simplified in TE than clusters [mk], [lg], [zb], [md₃] and [pʃ].

The predictions presented in (19) were first verified in a perception experiment. The results are presented in section 4.2.

4.1 Testing predictions of Hypothesis I for unattested clusters of English

Some of the “fillers” in the experiment presented in section 4.2 were unattested clusters of English and their simplified forms. These clusters were tested in the same experiment (in 4.0) and therefore under the same conditions as attested clusters (cf. Appendix 1, experiment 1). The setup is the same as presented before, except that the clusters of interest are now the unattested ones shown in (20). Recall that the task is discrimination

⁸ Although [lg] is unattested in English, it is attested in French and preserved in QF, a language that has the

of CVC₁C₂'# from CVC₁<C₂>#. For the unattested clusters, there were 12 sets of nonce syllables. As with the attested clusters in 4.0, there were 5 variants of the CV portion of the syllable. These were kept the same across all sets (cf. Appendix, experiment 1). There were also 2 presentations of each set. Each token had a dereleased C2.

- (20) a. CVmg', CVmk', CVmdz, CVm, CVlg', CVl
 b. CVzb', CVzd', CVz
 c. CVkp', CVkf, CVk
 d. CVpʃ, CVp

In (20a) the contribution of nucleus duration (mg-m vs. mk-m), high amplitude frication noise (mdz- m vs. mg-m), and transitions in /l/ (mg-m vs. lg-l) is being tested. In (20b), the contribution of the transition to the labial in /z/ is being tested. In (20c), the contribution of frication noise is being tested (kp-k and pʃ-p vs. kf-k).

4.2 Results

The d' values for discriminating unattested VC1C2#'s from VC1# are given in (21):

(21) d' values for distinguishing attested VC1C2#'s of English from VC1#.

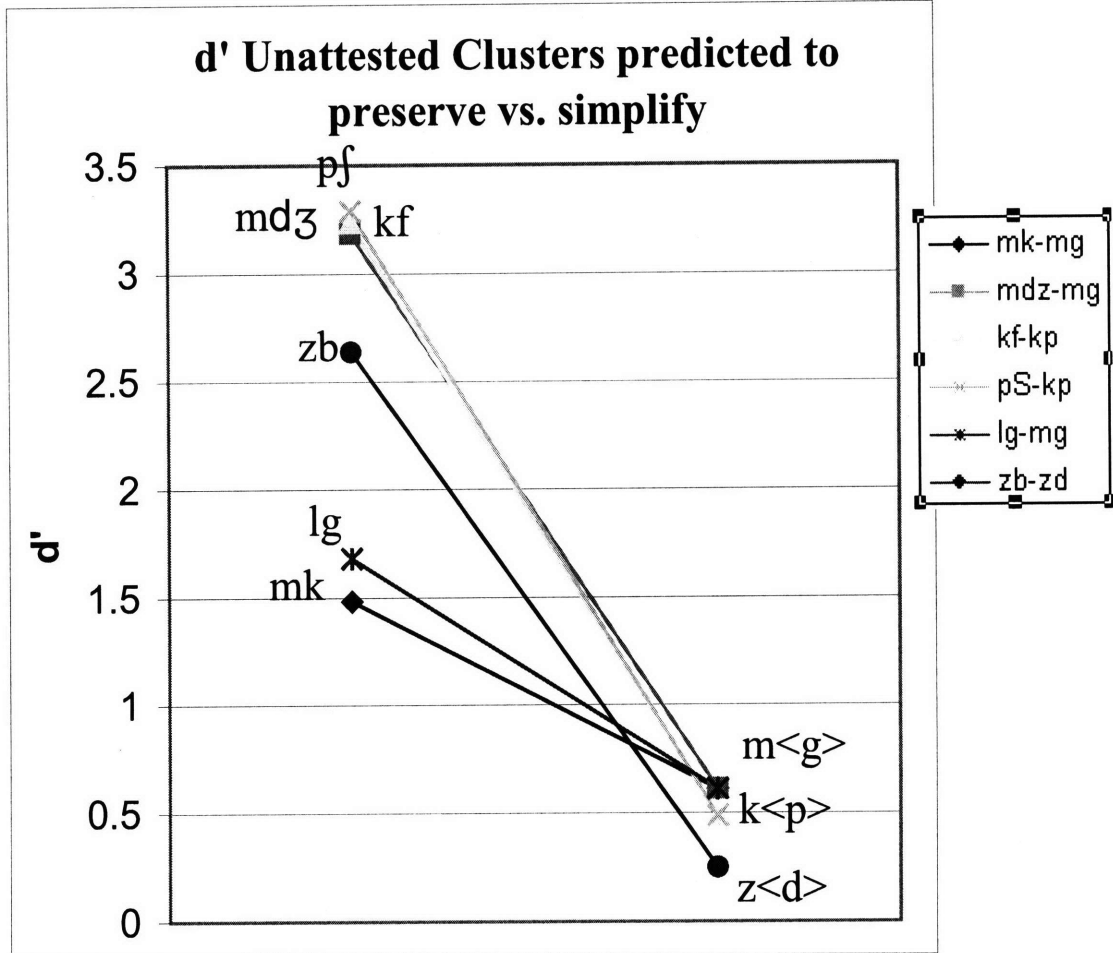
Unattested VC1C2#-VC1# Contrast	d'
[mg] – [m]	0.611949
[mk] – [m]	1.484788
[mdz] – [m]	3.185905
[lg] – [l]	1.682951
[zd] – [z]	0.250687
[kp] – [k]	0.486389
[kf] – [k]	3.237806
[pʃ] – [p]	3.289706

The d' values from (21) were contrasted for VC1C2's that are discriminable from VC1 versus VC1C2's that are not (p < .01 for each comparison).

same pattern of cluster simplification as TE, including preservation of [lb].

(22) d' values for two classes of final clusters with unreleased C2:

- (a) Clusters whose C2 lacks any of the internal or transitional cues listed in (18).
- (b) Clusters whose C2 has at least one of those cues.



The properties that are posited in Hypothesis I (section 2.0), for distinguishing VC1C2# from VC1#, are the only perceptually relevant ones: Release burst spectrum, values of F2 transitions to C2, transitions from /s/ to labial, nucleus (V+R) duration, and high amplitude frication noise.

We saw in (16) that C2 deletion in TE is triggered where one of these properties is not available to sufficiently distinguish attested C1C2's from C1. Now the rates of simplification of these unattested clusters will be measured to see if C2 deletion is a function of perceptual difficulty for distinguishing VC1C2# from VC1#.

5.0 TESTING HYPOTHESIS II WITH A PRODUCTION EXPERIMENT

In order to test the hypothesis that speakers encode the perceptually based difference between simplified and preserved clusters in their grammars, a production experiment was carried out. Participants were given attested and unattested clusters of English in nonce words, where VC1C2 was independently attested to be perceptually distinct or indistinct from VC1 in the previous experiment (cf. results in (16/17) and (21/22)). Judgments of grammaticality of VC1C2 were measured as rates of simplification of VC1C2 to VC1, in order to see whether or not TE speakers' judgments of cluster well formedness are indeed a function of perceptibility of C2, for both attested and unattested clusters.

5.1 Participants

The participants were 6 males and 5 females ranging in age from 21-52. Their ethnicities included white, latino, Afro-Trini, Indian, and Mixed. Seven participants were living in Northern Trinidad, mainly in the Trincity area and San Juan. One of the northerners was originally from Barrackpore. The other 4 live in the south: San Fernando, Saint James, and Couva.

5.2 Stimuli

TE subjects were presented with target stimuli consisting of nonce syllables X suffixed with -y, forming stimuli of the form X-y. Although cluster simplification in TE occurs word finally, it does not occur before the affix -y. Therefore, although a word like *frost* is simplified to /frɒs/ in a word-final context, the word *frosty* is realized as /frɒsti/.⁹ By adding the -y affix to stimuli, all word-final clusters of interest could be presented to participants intact, independently of whether or not those clusters would be simplified in a word final context.

⁹ A similar observation is made in Catalan (cf. Mascaro 1978). For example, clusters that simplify word finally do not simplify before the diminutive affix -et/.

The contents of X are shown in (23) (a) – (c). There were 20 sets of nonce –y affixed syllables, each corresponding to a different cluster. There were 5 variants of the CV portion of the syllable within a given set. These were kept constant across all sets.

(23) (a) X = [...VC₁C₂] sequences attested in English

(1) CVnt-y, CVnd-y, CVlt-y, CVld-y, CVlb-y

(2) CVsp-y, CVst-y

(3) CVpt-y, CVkt-y

(b) X = [...VC₁C₂] sequences unattested in English and shown to be discriminable from VC₁ even when C₂ is unreleased (Expt. 1)

(1) CVmd₃-y, CVmk-y, CVlg-y, CVlv-y

(2) CVzb-y, CVkf-y, CVp_f-y

(c) X = [...VC₁C₂] sequences unattested in English and shown to be indiscriminable from VC₁ when C₂ is unreleased (Expt. 1)

CVmg-y, CVzd-y, CVkp-y

The first type of X shown in (23a) had final clusters attested in English, some known to simplify in TE and some known to be preserved. The simplifying clusters (nd#, ld#, st#, pt#, kt#) were shown in (16/17) to be perceptually indiscriminable from their simplified forms in the absence of release. The other VC₁C₂'s in (23a) are preserved in TE, and were shown in (16/17) to be perceptually distinct from VC₁#. The second type of X shown in (23b) was comprised of tokens with VC₁C₂ sequences unattested in English, and shown to be perceptually discriminable from VC₁, when C₂ is unreleased (cf. 21/22). The third type of X shown in (23c) consisted of tokens with final VC₁C₂#'s unattested in English, and shown to be perceptually indiscriminable from VC₁# when C₂ is unreleased (cf. 21/22).

5.3 Procedure

X-y stimuli were presented to TE speakers visually on a 5x7 cue card, consisting of a sentence frame of the form given in (24), where X-y is a nonce adjective and X is a nonce verb. The cue cards that contained the stimuli and frame sentence were shuffled between participants so that the presentation would be randomized.

(24) They does call him¹⁰ X-y because he does ____ a lot. TE

Using examples without clusters as in (25), participants were trained to read the entire sentence given on the cue card, by substituting for the blank the nonce verb X. This would be achieved by stripping the adjective X-y of its affix.

(25) They does call him sleep-y because he does ____ a lot. TE

The blank must be filled in by X for the sentence to be grammatical since the habitual marker *does* takes the bare form of a verb.

(26) They does call him/he sleepy because he does sleep a lot.

‘They call him sleepy because he sleeps a lot’

By taking the X-y stimulus with the –y affix, and putting it in a grammatical context that requires removing the –y affix, all clusters in X could be read equally, in a context where simplification does not occur, while eliciting a response in a word-final context where simplification is possible. The purpose is to see what speakers do to clusters in the contexts where they can be simplified. After training, participants were free to fill in the blanks with any form that they deemed appropriate given affix removal. They were not told anything about simplifying clusters.

The well formedness of a cluster was measured as the rate at which TE speakers simplify a given cluster. Results reported here were coded in person by the experimenter in Trinidad, and recorded on a Marantz Professional portable solid state recorder PMD67, using a Crown GLM-100 microphone. Before presenting the results, some clarification is needed as to what counted as simplification and what did not for the coding of the data. Given the hypothesis that C2 deletes when VC1C2# is not sufficiently perceptually distinct from VC1#, one needs to control for the possibility that subjects produce full

¹⁰ In TE either “*him*” or “*he*” could be used in this context. Participants were encouraged to use whichever was most natural in their particular dialect of Trinidadian English.

CC#'s no matter what the cluster is, in all contexts, and that the rate of reported reduction is just the rate of cue-less, silent, C2 production.

In order to control for this possibility, the elicited nonce verb, X, was followed by a vowel (in “*a lot*”) so that if any C2 were pronounced it would have to be released, making it more audible to the experimenter. Participants were instructed not to pause before “*a lot*”. Later on, the experimenter isolated the target word and the modifier “*a lot*”, which came after the target word. As instructed during the experiment, participants did not pause between the tokens of interest and “*a lot*”. Hence stops produced before *a lot* were audibly released.

These isolated target words and “*a lot*” were numbered and organized into folders in Windows XP according to subject and cluster type. Each sound file was then coded by two second listeners (Zvi Bellin and Rachel Freedman) who are native speakers of American English. These second listeners were recruited from a Jewish cooperative house. They volunteered this service as a favor to the experimenter. Since target words and “*a lot*” were isolated, transcribers had no access to the target word with the affix –y, where simplification does not occur. For each subject, data for CVnt and CVnd were put into a folder labeled “nt.nd.n”. Data for CVlt, CVld, CVlb, CVlv, and CVlg were put into a folder labeled “lt.ld.lb.lg.lv.l”. Data for CVsp and CVst were put into a folder labeled “sp.st.s”. Data for CVpt and CVpf were put into a folder labeled “pt.p.psh”. Data for CVkt and CVkf were put into a folder labeled “kt.kf.k”. Data for CVmd₃, CVmk, and CVmg were put into a folder labeled “mj.mk.mg.m”. Data for CVzb and CVzd were put into a folder labeled “zb.zd.z”.

Transcribers were told that the label on each folder was indicative of what could come between the two vowels for each file in that particular folder. For example, in the “sp.st.s” folder for each subject, data could either be of the form CV_{sp}. *a lot*, CV_{st}. *a lot*, or CV_g. *a lot*. Second listeners were instructed to transcribe what they heard before “*a lot*”, but focusing on what came between the two vowels. Transcribers played the files in realPlayer audio by clicking on each file, and could listen to each file as many times as they wanted, in order to transcribe the data in that file. The transcriptions from the two second listeners matched the transcriptions of the experimenter. For no stimulus, for any given subject, was there disagreement in transcription between the experimenter and the

second listeners. Now that the procedure for coding has been clarified, results will be presented.

5.4 Results for Attested Clusters

Results with attested clusters in nonce words are consistent with the simplification pattern seen for real words in TE. The average simplification rate for each attested cluster is given in (27), for each subject, and averaged across subjects. For individual subjects, each figure is based on 5 presentations of the relevant cluster.

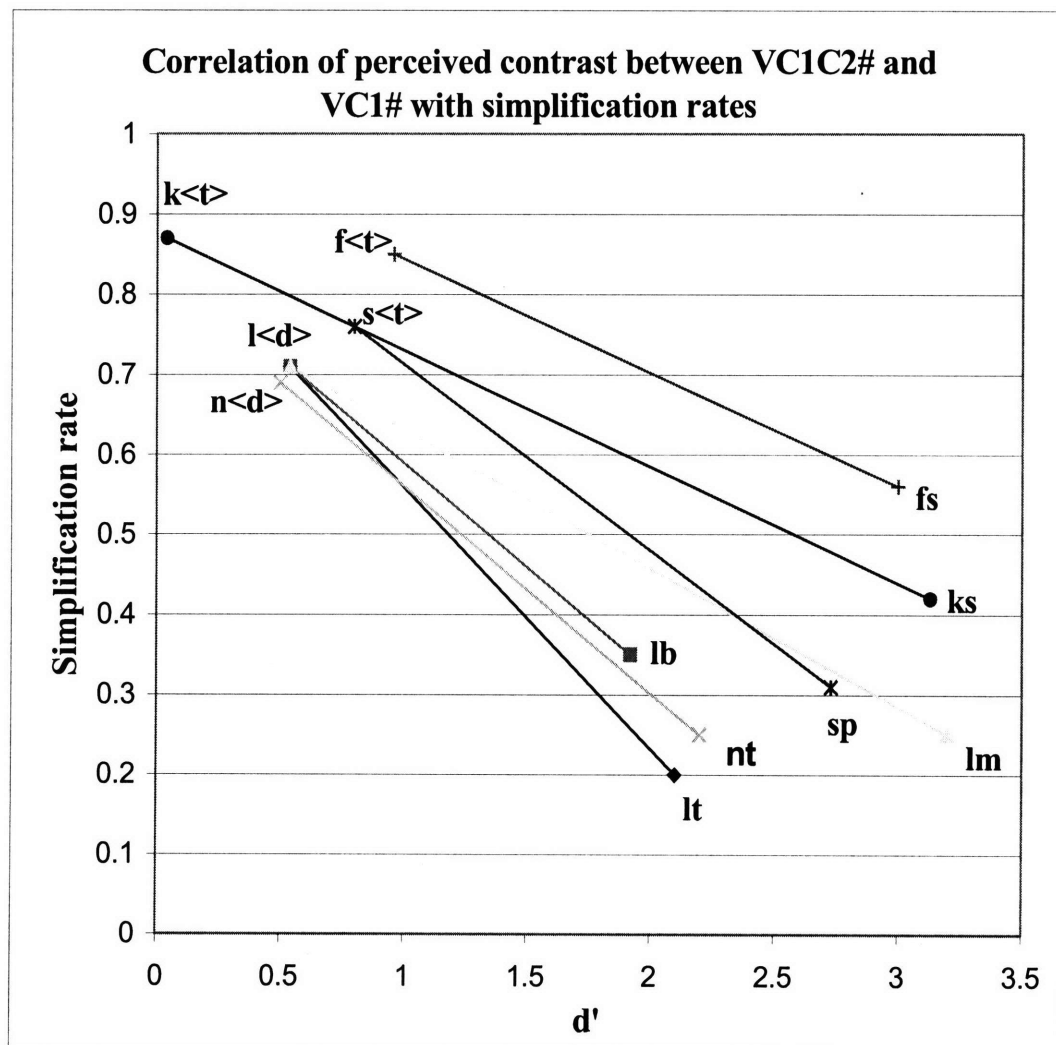
(27) Average simplification rates for attested clusters for 11 subjects (Subj)

Subj	[pt]	[kt]	[ks]	[ft]	[fs]	[st]	[sp]	[nd]	[nt]	[ld]	[lt]	[lb]	[lm]	[lv]
1	1	1	0.4	1	0	0.8	0.2	1	0.2	0.4	0	0	0	0
2	1	1	0.4	1	0.8	1	0	1	0.2	1	0	0	0	0.4
3	0.8	1	0.4	1	0.8	1	0.4	0.6	0.2	1	0.6	0.4	0.4	1
4	1	0.8	0.2	1	0.8	0.8	0	0.4	0.4	0.8	0	0.8	0.8	0.4
5	1	1	0.2	1	0.2	1	0.6	0.6	0.6	1	0.4	0.2	0	0.8
6	1	1	0.8	0.8	1	0.2	0.6	0.4	0.2	0.4	0.2	0.6	0.4	0
7	0.6	0.8	0.4	0.8	0.6	0.6	0	0.6	0	0.6	0	0	0	0
8	0	0.2	0.2	0	0.2	0.2	0	0	0	0	0	0	0	0.4
9	1	0.8	0.2	0.8	1	1	0.2	1	0.2	0.6	0	0.6	0.4	0.4
10	1	1	0.8	1	1	1	1	1	0.8	1	0.8	1	0.8	0.8
11	1	1	0.6	1	0.4	0.8	0.4	1	0	1	0.2	0.2	0	0
Avg	0.85	0.87	0.42	0.85	0.62	0.76	0.31	0.69	0.25	0.71	0.20	0.35	0.25	0.38

For each cluster, the average rate of simplification across all subjects, was examined as a function of d' for distinguishing VC1C2# from VC1# in the absence of release, as established in (16/17) and (21/22). For each token, every participant pronounced all C2's audibly before the -y affix.¹¹ However, in the same sentence frame, some C2's were produced when the -y affix was removed and some were not. Results here show that the rate of simplification is a function of perceptual difficulty in distinguishing VC1C2# from

VC1#. For each consonant cluster comparison, a nonparametric Wilcoxon signed-rank test was done in order to test for significant differences in simplification rates.¹² For every relevant minimal comparison of cluster types, a significantly lower d' for distinguishing that cluster from its simplified form, as established in (16/17), resulted in a higher rate of simplification as illustrated in (28). On the x-axis is d' , the measure of perceptual contrast. On the y-axis is the rate of simplification.

(28)



¹¹ When participants read the stimuli in the context before the affix – y, there was no simplification of VC1C2# strings since all C2's are cued in this word-internal, prevocalic context. Furthermore there was not deletion of C1 in VC1# since it is cued by the formants in the preceding vowel.

In general it is seen that there is significantly more simplification for clusters with a d' below 1 than clusters with a d' above 1.¹³

5.41 Simplification rates for [pt], [kt], [ks], [ft], and [fs] as a function of distinguishing C1C2# from C1#: Frication noise

There was a significantly higher rate of simplification for /kt/# than for /ks/# ($W+ = 55$, $W- = 0$, $N = 10$, $p \leq 0.001953$). There was also a significantly higher rate of simplification for /ft/# than for /fs/# ($W+ = 7.50$, $W- = 47.50$, $N = 10$, $p \leq 0.03711$). VC1C2's, /kt/# and /ft/#, are significantly less distinguishable from VC1# than ks# and fs# respectively (cf. 16/17). The simplification rate for /pt/# was also significantly higher than for /ks/# ($W+ = 4$, $W- = 62$, $N = 11$, $p \leq 0.006836$). Cluster /pt/# is significantly less perceptually distinguishable from /p/# than /ks/# is from /k/. The relevant property distinguishing VC1C2#'s with lower simplification rates from VC1C2#'s with higher simplification rates is frication noise in VC1C2# but not in VC1#.

5.42 Simplification rates for [st] versus [sp] as a function of distinguishing C1C2# from C1#: transitions in /s/

Cluster /st/# has a significantly higher simplification rate than /sp/# ($W+ = 2$, $W- = 53$, $N = 10$, $p \leq 0.005859$). Both /st/# and /sk/# are significantly less perceptually distinguishable from /s/ than /sp/# (cf. 16/17). The relevant property distinguishing /sp/, with its lower simplification rates, from /s/, is a transition from /s/ into the labial, which is not seen when /st/ is contrasted with /s/.

¹² This test was chosen since the distribution of differences between simplification for a given cluster comparison may be non-normally distributed, given that participants could respond as they wish after stripping the stimuli of their affixes.

5.43 Simplification rates for [nt] versus [nd] as a function of distinguishing C1C2# from C1#: nucleus duration

Cluster nd# had a significantly higher simplification rate than /nt/# ($W+ = 0$, $W- = 36$, $N = 8$, $p \leq 0.007812$). Cluster /nd/ has a significantly lower d' for perceptual distinctness from /n/ than /nt/# (cf. 16/17). The relevant property distinguishing /nt/# from /nd/# and /n/ is nucleus duration.

5.44 Simplification rates for [ld], [lt], [lb], [lm], and [lv] versus [ld] as a function of distinguishing C1C2# from C1#: nucleus duration, transitions in /l/

There was also a significantly higher rate of simplification for /ld/# than for /lt/#, /lb/#, /lm/#, and /lv/# (ld-lt: $W+ = 0$, $W- = 55$, $N = 10$, $p \leq 0.001953$; ld-lb: $W+ = 0$, $W- = 36$, $N = 8$, $p \leq 0.007812$; ld-lm: $W+ = 0$, $W- = 36$, $N = 8$, $p \leq 0.007812$; ld-lv: $W+ = 49.50$, $W- = 5.50$, $N = 10$, $p \leq 0.01953$). Cluster /ld/# is significantly less perceptually distinguishable from /l/ than lt#, lb#, lm#, or lv# (cf. 16/17). The relevant property distinguishing /lt/, with its lower simplification rates, from /l/, is nucleus duration, which is not seen when /ld/ is contrasted with /l/. The relevant property distinguishing /lb#, /lm/#, and /lv/# from /ld/# and /l/ is a transition in /l/, which is observed when /l/ forms a heterorganic cluster with C2.

5.5 Results for Unattested Clusters

For each unattested cluster, the rate of simplification was also examined as a function of d' for distinguishing VC1C2# from VC1#, in the absence of C2 release. The average simplification rate for each attested cluster is given in (29) for each subject, and averaged across subjects.

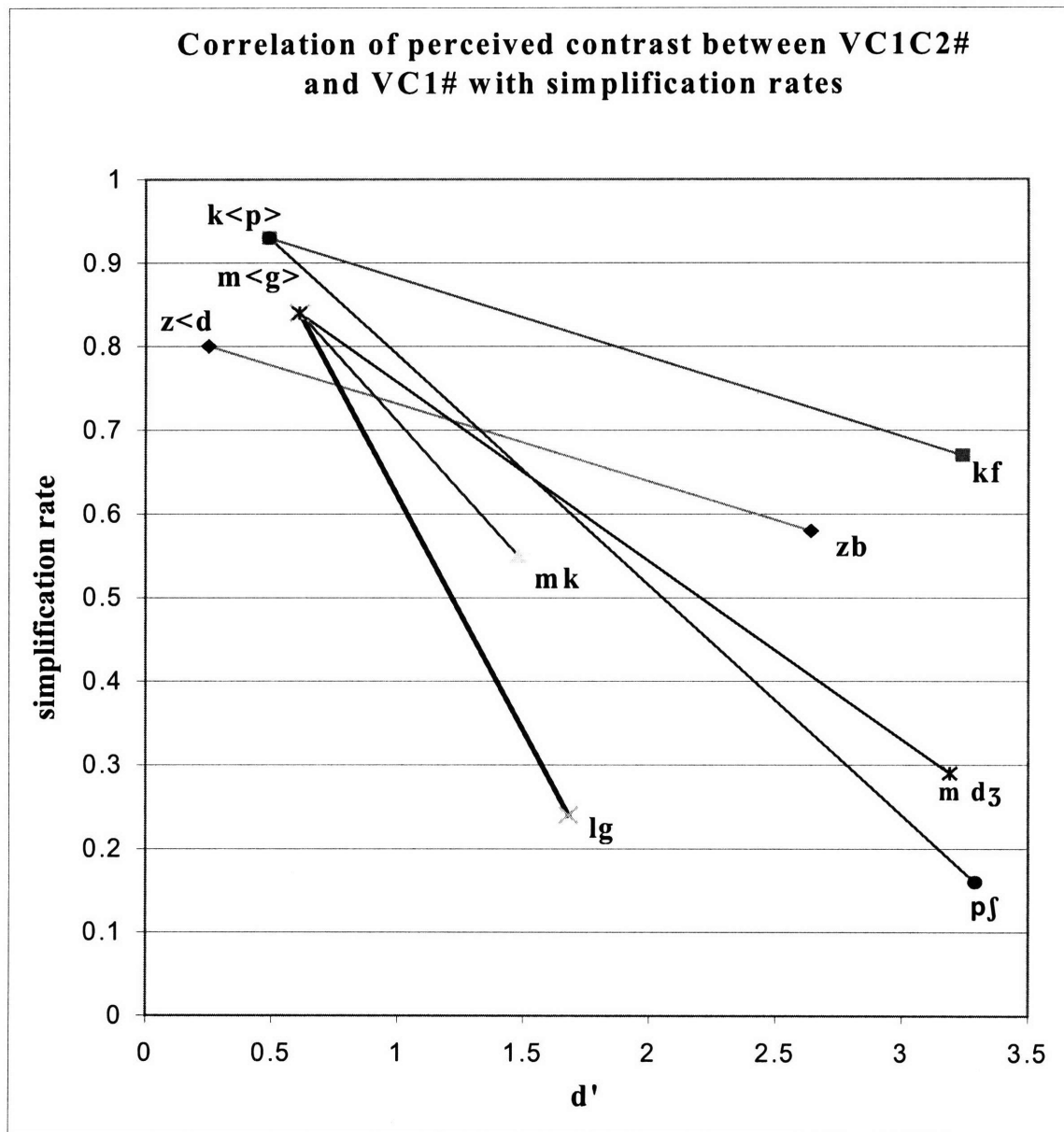
¹³ d' values are only useful as comparisons. However, using 1 as a reference point allows grouping of simplified versus preserved clusters.

(29) Average simplification rates for unattested clusters across 11 subjects (Subj)

Subj	[mg]	[mk]	[mdʒ]	[lg]	[kp]	[kf]	[pʃ]	[zd]	[zb]
1	0.80	0.40	0	0.20	1	0.8	0	0.8	0.6
2	1	0.20	0.40	0	1	0.4	0.2	0.8	0.2
3	1	0.20	0.40	0.20	1	1	0	1	1
4	1	0.20	0.80	0.80	1	1	1	1	1
5	1	1	0.40	0	1	1	0.2	1	1
6	1	1	0.40	0.20	1	1	0.2	1	1
7	0.40	0.40	0.20	0.40	0.40	0.40	0	0.2	0.2
8	0	0	0	0	0.8	0.20	0.2	0.2	0
9	1	0.80	0.40	0	1	0	0	1	0
10	1	0.80	0.20	0.8	1	0.80	0	1	0.8
11	1	1	0	0	1	0.80	0	0.80	0.6
Avg.	0.84	0.55	0.29	0.24	0.93	0.67	0.16	0.80	0.58

For every relevant minimal comparison of unattested cluster types, a significantly lower d' for distinguishing a VC1C2# from VC1#, as established in (16), resulted in a significantly higher rate of simplification. In general it is seen that there is significantly more simplification for clusters with a d' below 1 than clusters with a d' above 1. The same pattern follows for attested clusters.

(30)



**5.51 Simplification rates for [mg] versus [mk] as a function of distinguishing C1C2#
from C1#: nucleus duration**

There was a significantly higher rate of simplification for /mg/# than for /mk/# (: $W^+ = 0$, $W^- = 21$, $N = 6$, $p \leq 0.03125$). Cluster [mg]# has a significantly smaller d' for perceptual discrimination from /m/# than [mk]# (cf. 21/22). The relevant property distinguishing

/mk/#, with a lower simplification rate, from /mg/# with its higher simplification rate, is nucleus duration in VC1C2# but not in VC1#.

5.52 Simplification rates for [md₃] versus [mg] as a function of distinguishing

C1C2# from C1# & for [kf] and [pf] versus [kp] as a function of distinguishing

C1C2# from C1#: frication noise

Cluster /mg/# has a significantly higher simplification rate than /md₃/# ($W+ = 55$, $W- = 0$, $N = 10$, $p \leq 0.001953$). [mg] is significantly less perceptually distinguishable from /m/ than /md₃/# (cf. 21/22). The relevant property distinguishing /md₃/, with its lower simplification rates, from /m/, is frication noise, which is not seen when /mg/ is contrasted with /m/.

There was also a significantly higher rate of simplification for kp# than for kf# or pf# (kp-kf: $W+ = 21$, $W- = 0$, $N = 6$, $p \leq 0.03125$; kp-pf: $W+ = 0$, $W- = 55$, $N = 10$, $p \leq 0.001953$). Cluster /kp/# is significantly less distinguishable from VC1# than /kf/# or /pf/# (cf. 21/22). Although results in (21/22) indicate that the perception experiment was not sensitive enough to capture a perceptual difference between strident and non-strident fricatives, it is known that the latter is less perceptible than the former, due to lower amplitude frication noise (cf. Chomsky and Halle 1968, Miller and Nicely 1955). Although the simplification rate for /kp/# is higher than that of either /kf/# or /pf/#, we find that there was significantly more deletion for non-strident C2 of /kf/# than for strident C2 of /pf/# ($W+ = 45$, $W- = 0$, $N = 9$, $p \leq 0.003906$).

5.53 Simplification rates for [lg] versus [mg] as a function of distinguishing C1C2# from C1#: transitions in /l/

There was a significantly higher rate of simplification for /mg/# than for /lg/ ($W+ = 0$, $W- = 45$, $N = 9$, $p \leq 0.003906$). [mg] is significantly less perceptually distinguishable from /m/ than /lg/# is from /l/# (cf. 21/22). The relevant property distinguishing /lg/, with its

lower simplification rate, from /l/, is transitions in the /l, since C2 is heterorganic with /l/. Such transitions are not observable when /mg/ is contrasted with /m/.

5.54 Simplification rates for [zb] versus [zd] as a function of distinguishing C1C2# from C1#: transitions in /z/ as for /s/

Cluster /zd/# had a significantly higher simplification rate than /zb/# ($W^+ = 0$, $W^- = 21$, $N = 6$, $p \leq 0.03125$), the former having a significantly lower d' for perceptual distinctness from its simplified form than /zb/#. [zd] is significantly less perceptually distinguishable from /z/ than /zb/# (cf. 21/22). The relevant property distinguishing /zb/#, with a lower simplification rate, from /zd/# with a higher simplification rate, is a transition in C1 seen for VC1C2# but not for VC1# (cf. /sp/).

5.6 Comments on results

Hypothesis I (section 2.0) and Hypothesis II (section 4.0) are supported: for each relevant comparison, TE speakers have a higher rate of simplification for both attested and unattested clusters of English, where a VC1C2# is insufficiently perceptually distinct from VC1#. In general, both unattested and attested clusters with a d' below 1, simplify more than clusters with a d' above 1. Furthermore, the significantly higher simplification rate for C2 when it is a non-strident fricative in [kf] than for when C2 is a strident fricative, further corroborates the hypothesis that simplification occurs where VC1C2# is less perceptually contrastive with VC1#. This result underscores the importance of not just positing frication noise as a relevant perceptual dimension distinguishing VC1C2# from VC1#. A distinction must be made between high amplitude noise, as is characteristic of strident fricatives, and low amplitude noise as in non-strident fricatives. This is the reason why example (7) specifically posits high amplitude frication noise, and not just frication noise, as a relevant property providing sufficient perceptual contrast between VC1C2# and VC1#. In the next section I show how one should model a grammar where simplification is driven by perceptual difficulty, namely by the failure of VC1C2 to have at least one relevant property contrasting it with VC1#.

6.0 MODELING PERCEPTUAL FACTORS INTO THE GRAMMAR

Evidence has been presented here, that the grammar attributes simplification to the perceptual difficulty raised by unreleased C2's, by encoding perceptual differences between VC1C2's and VC1#. What would such a grammar look like? In this section I will sketch out a model showing what such a grammar would look like. We know the following:

- (31) (a) TE speakers don't release final stops.
(b) Unreleased C₂, in specific contexts, make it significantly harder to distinguish VC₁C₂'# from VC₁#, for any group of speakers.
(c) In exactly these contexts, and not others, the unreleased stops of TE are deleted

In order to account for these facts, we need (1) a component of phonetic realization that encodes cues like release and shortening as they would surface in each given language and (2) a component of the grammar that calculates the distinctiveness between C₁C₂#-C₁#, and neutralizes them as a consequence of subminimal contrast. The components of such a grammar can be found in Flemming (2006). Flemming's model can be applied to coda cluster simplification. This model has 3 components. The inventory component selects a basic inventory of maximally contrasting segment types. The hypothetical realization component maps strings of segments drawn from the inventory, onto a phonetically realized representation. This component projects how a string of speech sounds would be realized in a given language. This mapping is done via articulatory constraints. The third component *evaluates surface contrasts*. Distinctiveness of candidate contrasts is evaluated by perception-based constraints that neutralize those contrasts that are not sufficiently distinct. I will now illustrate with some examples how such a model can be applied to coda cluster simplification

Since the *inventory* component does not play a role in the current discussion, the reader is referred to Flemming (2004, 2006) for more detail. For the current discussion it is sufficient to know that the constraints of the inventory generate inventories of phonemes that are maximally contrastive with each other. For example, it is known that

the release burst spectrum of any given stop consonant makes that stop more contrastive with other stops (cf. Dorman, Studdert-Kennedy and Raphael 1977; Kewley-Port et al. 1983; Lisker 1999; Wright 2004). Therefore the inventory generates all stops with their characteristic release burst spectrums.

I will make Flemming's model more explicit by postulating that in all components, acoustic information is encoded for each segment of speech. This is analogous to distinctive features being part of the representation of each phoneme. Namely, perceptual dimensions are included as part of the representation of segments of speech. For example, the representation of a stop /k/, from the inventory, includes the release burst spectrum of that stop as generated in the inventory. The representation for the stop /k/, in the grammar, would not only include features such as [+velar] and [-voice]. The representation of /k/ also reflects the acoustic information for /k/, namely that /k/, as generated in the inventory, has its characteristic release burst spectrum. This is indicated by the notation {RB_k}.

(32) Representation of /k/ with its release burst spectrum. ^l = release, RB = Release burst

k ^l
{RB _k }

Such encoding of acoustic information allows the motor system to produce sounds that are characteristic of a given language.

The second component of Flemming's model generates a *Hypothetical Surface Realization* that is characterized by articulatory constraints that map the segments of the inventory onto their phonetic realization. For the current analysis, a hypothetical realization is hypothetical in the sense that it specifies the realization a cluster would have, if it were permitted in TE. I will illustrate how this model works with a simplified TT cluster. Let us take for example the strings /æk/ and /ækt/ in TE. We know that TE neutralizes /Vkt/ and /Vk/ to /Vk/. For example *pact* and *pack* are neutralized to /pæk/ in TE. The input of segments /æ/ and /k/ to the *Hypothetical Surface Realization* component, from the inventory, is represented as follows for the string /æk/:

(33) RB = release burst; $Tr_{\text{æ-k}}$ = Transitions from /æ/ into /k/

æ_1	k_2^t
	$\{\text{RB}_k\}$
	$\{Tr_{\text{æ-k}}\}$

Since all stops that are output from the inventory have release burst spectrums, these release bursts (RB) are part of the representation for the /k/. The transitions from /æ/ into /k/ must also be part of the representation of /æk/, since perceptual dimensions, including formant transitions, are included as part of the representation of segments of speech.

For the token /ækt/, the output of segments /k/ and /t/, from the inventory, into the *Hypothetical Surface Realization* component, is characterized by release of both /k/ and /t/. The grammar encodes perceptual dimensions for each segment of the string /æk^t/. Therefore the release bursts of /k/ and /t/, and the transitions from /æ/ into /k/ are part of the representation of /ækt/, when segments are input from the inventory and merged as strings in the *Hypothetical Surface Realization* component.

(34) RB = release burst; $Tr_{\text{æ-k}}$ = Transitions from /æ/ into /k/

	æ_1	k_2^t	t_3^t
		$\{\text{RB}_k\}$	$\{\text{RB}_t\}$
		$\{Tr_{\text{æ-k}}\}$	

Perceptual dimensions like transitions from /æ/ into /k/, that are introduced when segments from the inventory are merged, as well as perceptual dimensions like release that are generated in the inventory, are part of the representation of strings as in (33) and (34). These strings of segments are then evaluated by the *Hypothetical Surface Realization* component. In Trinidad English, this component has a ranking of articulatory constraints that act upon the representations in (33) and (34), /æk^t/ and /æk^tˈt/, in order to generate a hypothetical surface form that represents these strings as they would be realized in TE. At this level of evaluation, acoustic parameters from the inventory can be changed by constraint rankings that are language specific.

The correct hypothetical realizations of the strings, /æk^t/ and /æk^tˈt/, as they would be generated in TE, is given in (35). C1 and C2 are not released in TE. There is also shortening of the vowel before a voiceless stop in TE.

(35) \check{v} = short vowel

(a) Inventory: /æ^k/ → Hypothetical Surface Realization for TE: /æ^k/

(b) Inventory: /æ^kt/ → Hypothetical Surface Realization for TE: /æ^kt/

The encoding of perceptual dimensions in the representations of these outputs of the *Hypothetical Surface Realization* component is as follows:

(36)

(a) RB = release burst; $Tr_{\text{æ-k}}$ = Transitions from /æ/ into /k/

æ ₁	k ₂
	{ $Tr_{\text{æ-k}}$ }
	\check{v}

(b) RB = release burst; $Tr_{\text{æ-k}}$ = Transitions from /æ/ into /k/

æ ₁	k ₂	t ₃
	{ $Tr_{\text{æ-k}}$ }	
	\check{v}	

There are no release bursts for C1 or C2, and there is shortening of the vowel, and transitions from the vowel into C1. The lack of release, and shortening of the vowel must be generated by articulatory constraints of the *Hypothetical Surface Realization* component. These constraints will not be given here. The main point is that both the loss of the perceptual dimension of release, and the generation of the perceptual dimension of vowel duration in TE, must be reflected in the representations of the outputs of the *Hypothetical Surface Realization* component, as shown for /æ^kt/ and /æ^k/ in (36).

The third component *evaluates surface contrasts* for the outputs of the *Hypothetical Surface Realization* component. The outputs of (36) (a) and (b), with the encoding of an inventory of properties that can distinguish VC1C2 from VC1, reveals that there is no contrast between /æ^k/ and /æ^kt/ along the relevant parameters. VC1C2# and VC1# are neutralized when there are no relevant acoustic properties sufficiently distinguishing VC1C2# from VC1#. Both /æ^k/ and /æ^kt/ in TE have formant transitions from /æ/ into /k/. Recall that the vowel before the voiceless stop has comparable duration in the two cases, and does not differ in any properties that are relevant for sufficient contrast. It is indeed the case that *pact* and *pack* are neutralized to *pack* in TE.

Having given an overview of a model that can account for the facts in the current paper (cf. 30), I will now go into more detail on how the third component, *Evaluation of surface contrasts*, works. This will be illustrated with the hypothetical strings “amg” and “am” in (37). The hypothetical realization of these strings in TE should encode the fact that C1, /m/, has formant transitions from the vowel and that C2 in *amg* is unreleased. Example (37) is a schematic representation of the perceptual differences characterizing the contrast between [æmg] and [æm].

(37) TE outputs of *Inventory* and *Hypothetical Surface Realization* component, with encoding of relevant properties perceptually distinguishing VC1C2# from VC1#

ǃ = short vowel, Tr = Transitions, Ø = null segment

	(a) æ ₁	m ₂	g ₃
Relevant properties		{Tr} _{a-m2}	
	(b) æ ₁	m ₂	Ø ₃
Relevant properties		{Tr} _{a-m2}	

The representations in (37) illustrate the point that, for a hypothetical string like “am” in TE, the inventory of relevant cues is the same as for “amg”: transitions into segment 2, /m/, in the vowel.

For a hypothetical string like “amk”, the grammar does not only keep account of transitions into /m/ from the vowel, / æ /, but also the vowel shortening before /k/. This distinguishes [æmk] from [æm], as shown in (38). This nucleus shortening property is a predictable consequence of articulatory constraints. The result is that “amk” is more perceptually distinct from “am” than “amg” is from “am”.

(38) TE outputs of *Inventory* and *Hypothetical Surface Realization* component, with encoding of relevant properties perceptually distinguishing VC1C2# from VC1#

ǃ = short vowel, Tr = Transitions, Ø = null segment

	(a) æ ₁	m ₂	k ₃
Relevant properties	ǃ from k ₃	{Tr} _{a-m2}	
	(b) æ ₁	m ₂	Ø ₃
Relevant properties		{Tr} _{a-m2}	

We have seen thus far that the grammar outlined above generates representations for the pair {amk, am}, which are distinguished by nucleus duration. The same grammar

generates representations for the pair {amg, am}, which are not distinguished by any of the relevant auditory properties discussed earlier.

These hypothetical strings (realizations of VC1C2 and VC1) from (37) and (38), generated in the *Hypothetical Surface Realization* component, are the input to the next component, “*Evaluation of Surface Contrasts*”, which assesses distinctiveness. Here I will propose a unique constraint –*MinimalContrast*- which evaluates the distinctiveness of pairs of strings like VC1C2 and VC1.

(39) *MinimalContrast* (xy-ab; transition-L, transition-s, noise, release burst, vowel duration): A contrast between the strings xy and ab is permitted in context *k* only if the acoustic difference between xy/ __*k* and ab/ __*k* is non null.

*Where x, y, a, b = segments, and
For x, y: x, or y, or both can be the null string, and
For a, b: a, or b, or both can be the null string
Voicing during closure is not a relevant cue*

Given 2 strings xy and ab where x, y, a, and b can be segments or the null string, a contrast between these 2 strings xy and ab are permitted in context *k*, only if the acoustic difference between strings xy and ab in context *k* is non-null as defined by the relevant properties. The relevant properties are transitions in L, transition in /s/, high amplitude frication noise, release bursts, and vowel shortening. For the strings “amg” and “am”, the inventory of relevant properties is the same as shown in (37) and (40). For “amg”, *x* = /m/ and *y* = /g/. For “am”, *a* = /m/ and *b* is null.

(40) *Evaluation of Surface Contrasts* between VC1C2# (Vmg#) and VC1# (Vm#)

Tr = Transitions, Ø = null segment

	(a) æ_1	$m_2 = x$	$g_3 = y$	Contrast xy-ab
Relevant properties		{Tr} _{a-m2}		
	(b) æ_1	$m_2 = a$	Ø ₃ = b	
Relevant properties		{Tr} _{a-m2}		
C1C2 – C1	0	0	0	0 relevant cues

There is zero contrast between “amg” and “am”, as indicated by a difference of zero relevant properties. MinContrast is violated. Isolated or phrase final forms must be

privileged in the grammar for computing perceptual distinctness between VC1C2 and VC1, since stops are inevitably released in prevocalic contexts.

In contrast with “amg”, the inventory of cues for “amk” contrasts with that of “am” by one relevant cue (vowel shortening “ \check{v} ”) as shown in (38) and (41).

(41) *Evaluation of Surface Contrasts between VC1C2# (Vmk#) and VC1# (Vm#)*

Tr = Transitions, \check{v} = vowel shortening, \emptyset = null segment

	(a) æ_1	$m_2 = x$	$K'_3 = y$	Contrast xy-ab
Relevant properties	\check{v} from k'_3	$\{\text{Tr}\}_{a-m_2}$		
	(b) æ_1	$m_2 = a$	$\emptyset_3 = b$	
Relevant properties		$\{\text{Tr}\}_{a-m_2}$		
C1C2 – C1	(1) \check{v} from k'_3	0		1 relevant cue

MinContrast only needs one cue to be satisfied. Therefore the contrast between “amk” and “am” does not violate it, while the contrast between “amg” and “am” does. Example (44) illustrates that the candidates for MinContrast are inventories of strings. When the relative well formedness of “amg” versus “amk” is evaluated in an inventory that includes “am”, MinContrast is violated for “amg” as shown in (44) but not for “amk” as shown in (45). A constraint **Merge* penalizes merging of potential contrasts, and **Gesture* penalizes each gesture.

(42) **Merge*: No Mergers of input forms allowed, i.e no neutralization.

(43) **Gesture*: No articulatory gestures allowed.

**Merge* is ranked below *MinContrast* in order for neutralization to occur. **Gesture*, which penalizes each gesture, results in neutralization to VC1 and not VC1C2. We will see in (45) that **Gesture* must be ranked below **Merge*

(44) Evaluation of Inventories of TE VC1C2 and VC1 by MinContrast
MinContrast >> *Merge

/æ ₁ m ₂ g ₃ /a- /æ ₁ m ₂ /b	Min Contrast	*Merge	*Gesture
(a) /æ ₁ m ₂ g ₃ /a, /æ ₁ m ₂ /b	*!		***
(b) /æ ₁ m ₂ g ₃ /a,b		*	***!
(c) \varnothing /æ ₁ m ₂ /a,b		*	**

For a non-releasing language like TE, the winner is an inventory where “amg” and “am” are neutralized.

The contrast between “amk” and “am” is better than the contrast between “amg” and “am” as shown: there is one relevant property, vowel duration, contrasting “amk” with “am”. MinContrast is ranked above *Merge, which in turn is ranked above *Gesture.

(45) Evaluation of Inventories of TE VC1C2 and VC1 by MinContrast
*Merge >> *Gesture

/æ ₁ m ₂ k ₃ /a- /æ ₁ m ₂ /b	Min Contrast	*Merge	*Gesture
(b) \varnothing /æ ₁ m ₂ k ₃ /a, /æ ₁ m ₂ /b			***
(b) /æ ₁ m ₂ k ₃ /a,b		*!	***
(c) /æ ₁ m ₂ /a,b		*!	**

Therefore, we expect neutralization of “amg” and “am” but not “amk” and “am”. This correlates with the higher rate of simplification that was observed for /mg/# over /mk/# in experiment 2 (cf. 28). The type of analysis proposed here derives the pattern of simplification for real words of Trinidad English. It also accounts for the contrasts in simplification rates found for unattested clusters with VC1C2’s that are perceptually discriminable from CV1, versus C1C2’s that are not. The current analysis also captures the crucial role that the language-specific generation of cues like release plays in cluster simplification.

Example (46) shows how MinContrast would work in a language like SE or SF where C2 is typically released. Both “amg” and “amk” would satisfy MinContrast, since both have a release burst contrasting VC1C2# from VC1#. The string “amk” also has the vowel shortening cue.

(46) (a) Evaluation of Inventories of Standard English and Standard French VC1C2 and VC1 by MinContrast. Tr = Transitions, RB = Release burst, Ø = null segment

	(a) æ ₁	m ₂ = x	G ₃ ^l = y	Contrast xy-ab
Relevant properties		{Tr} _{a-m2}		
	(b) æ ₁	m ₂ = a	Ø ₃ = b	
Relevant properties		{Tr} _{a-m2}		
C1C2 – C1	0	0	(1) RB	1 relevant property

(b) Evaluation of Inventories of Standard English and Standard French VC1C2 and VC1 by MinContrast. ˇ = V shortening, Tr = Transitions, RB = Release burst, Ø = null segment

	(a) æ ₁	m ₂ = x	K ₃ ^l = y	Contrast xy-ab
Relevant properties	ˇ from k ₃	{Tr} _{a-m2}		
	(b) æ ₁	m ₂ = a	Ø ₃ = b	
Relevant properties		{Tr} _{a-m2}		
C1C2 – C1	(1) ˇ from k ₃		(2) RB	2 relevant properties

Having at least one relevant cue distinguishing a cluster from its simplified form satisfies MinContrast. This accounts for how it is that the subset of clusters that are simplified in non-releasing languages like TE and QF are preserved in releasing languages like SE and SF.¹⁴

7.0 WORD-FINAL CODA CLUSTER SIMPLIFICATION CROSSLINGUISTICALLY

In the current model, variation in the typology of cluster simplification comes from the ranking of articulatory constraints in any given language, for relevant properties like release of C2 and vowel shortening before a voiceless stop. Non-C2-releasing languages like QF, Cameroon English (CE), African American English, and Jamaican Creole have exactly the same pattern of simplification reported for TE in (1) – (5) (cf. Bobda 1994 on

¹⁴ Although some Standard varieties of English release final consonants variably, there is consistent release of C₂ in clusters compared to singleton final stops (Kang 2003, Albright 2006).

CE; Green 1992 on AAE; Akers 1981, Meade 2001, Patrick 1991, and others on Jamaican Creole).¹⁵ Data will now be shown for QF in (47). These judgments were elicited from two native speakers of QF.

(47) Patterns of C-deletion in QFE: Sonorant-stop coda clusters

(a) Son = Sonorant; T = voiceless stop; F = fricative (voiced/voiceless);

<...> = deleted Consonant; C' = unreleased stop

	(1) VSon ₁ T ₂ #	(2) Vson ₁ d ₂ #	(3) VL ₁ B ₂ #
SF	[sɛ̃nt] “cent” [æ̃sɪlt] “insult”	[bæ̃nd] ¹⁶ “bande” [sɔld] “solde”	[bɪlb] “bulbe” [ælg] “algie”
QF	[sɛ̃ntʰ] “cent” [koltʰ] “colt”	[bæ̃n] <d> “bande” [sɔl] <d> “solde”	[bʌlbʰ] “bulbe” [ælgʰ] “bulbe”

(b) Patterns of C-deletion in QF: Obstruent-obstruent coda clusters

	(1) VT ₁ T ₂ #	(2) VF ₁ t ₂ #	(3) Vs ₁ p ₂ #
SF	[pækt] “pacte” [æksɛpt] “accepte”	[lɪft] “lift” [rɛst] “reste” [rɪsk] “risque”	[krɪsp] “crispe” [lɪsp] “lispe”
QF	[pækʰ] [æksɛpʰ]	[lɪfʰ] “lift” [rɛsʰ] “reste” [rɪsʰ] “risque”	[krɪspʰ] “crispe” [lɪspʰ] “lispe”

These data show that the following factors also play a role in the preservation of clusters in QF: voicing, heterorganicity, labiality, and frication noise. Given that the QF simplification pattern is the same as for TE, the current analysis of simplification applies straightforwardly to QF. Deletion of C₂ is triggered when the distinctiveness of VC₁C₂# and VC₁# falls below a particular threshold. The relevant cues are the release burst spectrum, values of F2 transitions to C₂, transitions from /s/ to labial, nucleus (V+R) duration, and high amplitude frication noise.

For languages with the same pattern of simplification as TE, we expect that the status of release and vowel shortening is the same. These factors vary crosslinguistically. As

¹⁵ Similar patterns of simplification have also been reported for some dialects of Dutch (cf. Goeman 1999; Hinskens 1992, 1996; Shouten 1982, 1984).

¹⁶ These are loan words of English therefore there is some realization of the nasal and not just a nasalized vowel.

stated previously, QF does not release word-final consonants (cf. Archambault & Dumochel 1993). It has also been verified that QF has nucleus shortening before a voiceless consonant.

Nonce tokens were recorded in isolation from 2 native speakers of QF. For each speaker, each token was recorded 5 times. The average duration was computed. Simplification is optional in QF, so the effect of C2 on nucleus duration was measured for all clusters in a word-final context. The use of a double “n” in the orthography in (48) resulted in QF speakers producing a nasal after the vowel and not just nasalizing the vowel. For tokens where C1 is a sonorant, duration measurements are of the preceding vowel and this sonorant. For tokens where C1 is a stop, duration measurements are of the preceding vowel.

(48) Average nucleus duration in QF as a function of voicing of C2

VC ₁ C ₂ - VC ₁	Δ (in ms)	C2 status in QF
VND-VN	2	Reduced
VNT-VN	95	Preserved
VLD-VL	4	Reduced
VLT-VL	110	Preserved
VTT-VT	11	Reduced

Stimulus list: VND: pannde; VNT: pannte, pannque; VN: panne; VLD: palde; VLT: palte, palque; VN: palle VTT: pafte, papte, pakte; VT: paffe, pappe, paque.

The degree of shortening for clusters with voiceless C2 (NT and LT) is significantly greater than for clusters with voiced C2 (ND and LD). Furthermore, there is no cumulative shortening. A vowel followed by two voiceless obstruents is as short as one followed by a single voiceless obstruent. Therefore the duration of the vowel does not distinguish Vkt from Vk or Vst from Vs.

For a language like Catalan, that does not have vowel shortening or word-final stop release (Mascaro p.c, Wheeler 1979), we expect that the pattern of simplification is the same, except that C2’s that are cued in TE by the duration of the reduced vowel, would also be simplified in Catalan. This is indeed the case. The pattern of simplification in

Catalan is pretty much the same as for TE, except that among the sonorant-voiceless stop clusters, only heterorganic clusters are preserved. The heterorganic sonorant-voiceless stop clusters of Catalan include [lk] and [lp]. Even in the absence of nucleus shortening before voiceless stops in a language like Catalan, these are expected to be preserved since they are heterorganic like preserved [lb] of TE and QF, and preserved [lg] of QF. [lk] and [lp] are predicted to be sufficiently perceptually contrastive with /l/, due to values of F2 transitions to C2 in /l/.

For languages with release of C2 we expect, as mentioned previously, that the subset of VC1C2's that have no cues sufficiently contrasting them with VC1 in the absence of release, would be sufficiently contrastive with VC1. This subset includes [ld], [nd], [st], [sk], and [TT]. It is indeed the case than when non-releasing languages like TE and QF are compared to their C2 releasing, SE and QF counterparts (cf. Albright 2006, Kang 2003 on English), this subset of clusters are simplified in the former and not in the latter.

Further acoustic studies are needed to examine simplification patterns in light of release and vowel shortening crosslinguistically. If the release status of C2 and nucleus shortening do not give a complete typology of simplification in all languages as they do for example in SF, QF, and Catalan, then more variation can be introduced into the current system. Different MinContrast constraints can be introduced into the system that differ as to what are the relevant properties distinguishing VC1C2 from VC1. This would predict that any one or all of the groups of clusters that are preserved, due to any one cue or a combination of cues in one language, can be simplified in another language, if none of those cues are relevant in that other language. Variation due to different relevant cues could be constrained if a crosslinguistically constant hierarchy of relevant cues exists. The complete set of possibly relevant cues would be the ones that are proposed here: release burst spectrum, values of F2 transitions to C2, transitions from /s/ to labial, nucleus (V+R) duration, and high amplitude frication noise. The data in languages explored or mentioned here, however, do not necessitate adopting different relevant cues across different languages as a source of variation in cluster simplification patterns.

Either source of variation (based on rankings of articulatory constraints for nucleus shortening or of MinContrast with differences in relevant cues) is certainly more constrained than an analysis like Côté's (2004). Côté's account predicts many

simplification patterns but does not account for the pattern of simplification reported here for TE and QF. This will be discussed further in the next section. Her analysis is certainly on the right track since it is a perceptual approach. However, with seventeen constraints, it predicts 355,687,428,096,000 logically possible grammars and 787 different simplification patterns. This was verified by the OTSoft software (Hayes, Tesar and Zuraw 2003). In fact, however, the pattern of cluster simplification is rather consistent crosslinguistically. The current analysis captures this.

8.0 DISCUSSION/CONCLUSION

This paper has demonstrated a previously unreported correlation between the languages that do not release C2 and cluster simplification, as well as a striking correlation between the availability of specific auditory properties (including release), and the set of C2's that are deleted word finally across languages.

A further finding has been that perception plays a role in the computation of grammatical well-formedness. Perceptual differences between VC1C2# and VC1# are encoded in the grammar. The relevant properties that perceptually distinguish VC1C2# from VC1# are the release burst spectrum, values of F2 transitions to C1 or C2, transitions from /s/ to a labial, Nucleus (V+R) duration, and high amplitude frication noise. Speakers encode an inventory of these relevant properties and delete C2's where there is subminimal contrast between VC1C2# and VC1#, not only for attested strings of their language, but also for unattested strings of their language. This shows that the perceptual phenomenon explored here is not just a historical process where some C₂'s were lost over time, resulting in a grammar with a set of constraints that simply characterizes the set of attested clusters in terms of features, as claimed by some analyses (cf. Côté 2004). There is evidence for a component of the phonological grammar where speakers have and process knowledge of how hypothetical strings of their language would be phonetically realized in terms of perceptual dimensions. This encoding in turn can be used by perceptual constraints like the MinContrast constraint proposed here, where subminimal contrasts in a language can be neutralized.

Earlier work (e.g. Côté 2004) proposes a correlation between cluster simplification and similarity between C_1 and C_2 in terms of features, based on the idea that the more similar C_2 is to C_1 , the less perceptible C_2 is. By referring to similarity between C_1 and C_2 in terms such as homorganicity, and not referring directly to specific properties that perceptually distinguish VC1C2 from VC1, this approach fails to account for all the data. Côté's account does not predict that TE and QF preserve /sp/# but not /sk/#. C_1 and C_2 for both of these clusters are equally similar. Furthermore, such an analysis does not predict that TE speakers preserve unattested /lg/# but not /mg/#. Again C_1 and C_2 are equally similar in [mg] and [lg]. Furthermore, Côté's account does not explain why TE speakers have a higher rate of preservation for /mdʒ/# and /pʃ/ versus /kf/#. The current account draws the right distinction between strident and non-strident frication noise.

That TE speakers do not simplify unattested clusters across the board, but are more likely to preserve nonce clusters to the extent that C_2 is more perceptible, shows that these speakers know about perceptibility as modeled in section 6. However, two questions remain unanswered in the current study and will be taken up in further research. The graphs in (27) and (29) show that simplification rates increase as the perceptual difference between a cluster and its simplified form decrease, as measured by d' values. These graphs also suggest inconclusively that among clusters that are more likely to be preserved because of higher d' values, simplification rates are actually higher for clusters with more perceptible C_2 's.¹⁷ Why is this the case? A line of best fit among these preserved clusters on the right side of these graphs looks as though it would have a positive slope. If this observation is valid, then this issue must be addressed. This indicates that even though simplification of VC1C2 is a function of perceptual contrast between VC1C2 and VC1, there might be some competing, unidentified factor that favors deletion of C_2 's that are more perceptible. A further question is left over. Perception is one dimension of well-formedness but not the only one. For unattested clusters, why are subjects willing to produce clusters that might be otherwise ungrammatical provided that C_2 is perceptible? The striking discovery that perceptual differences between VC1C2# and VC1# are encoded in the grammar, and that a VC1C2# surfaces only if there is at least one relevant cue contrasting it with VC1#, raises these provoking questions.

¹⁷ Thanks to Pat Keating for pointing this out.

Appendix

Experiment 1

(a) Target stimuli (attested clusters and simplified forms)

gamp, gam, emp, em, zimp, zim, zump, zum, deemp; yant, yan, unt, un, zint, zinn, yunt, yun, zeent, zeen; yand, yan, und, un, zind, zinn, yund, yun, zeend, zeen; dakt, dack, ekt, eck, zikt, zick, zukt, zuck, deekt, deek; daks, dack, eks, eck, ziks, zick, zuks, zuck, deeks, deek; dalb, dal, ulb, ull, zilb, zill, yulb, yull, beelb, beel; dalt, dal, zilt, zill, yult, yull, ult, ull, beelt, beel; dald, dal, uld, ul, zild, zill, yuld, yull, beeld, beel; dalk, dal, zilk, zill, yulk, yull, ulk, ull, beelk, beel; dalm, dal, ulm, ul, zilm, zill, yulm, yull, beelm, beel; ; dalv, dal, ulv, ul, zilv, zill, yulv, yull, beelv, beel; dasp, dass, esp, ess, esp, zisp, ziss, yusp, yuss, deesp, dees; esk, ess, zisk, ziss, yusk, yuss, deesk, dees; dast, dass, est, ess, zist, ziss, yus, yuss, deest, ees, dask, dass; daft, daff, eft, eff, zift, ziff, zuf, zuff, deeft, deef; dafs, daff, efs, eff, zifs, ziff, zufs, zuff, deefs, deef.

(b) “Fillers” (target stimuli for unattested clusters and simplified forms)

gamk, gam, emk, em, zimk, zim, zumk, zum, deemk, deem; gamge, gam, emge, em, zimge, zim, zumge, zum, deemge, deem; gang, gam, emg, em, zimg, zim, zumg, zum, deemg, deem; dalg, dal, ulg, ul, zilg, zill, yulg, yull, beelg, beel; dakf, dack, ekf, eck, zikf, zick, zukf, zuck, deekf, deek; dakp, dack, ekp, eck, zikp, zick, deekp, deek; dazb, daz, ez, zizb, zizz, zuzb, zuzz, deezb, deez; dazd, daz, ez, zizd, zizz, zuzd, zuzz, deezd, deez; vapsh, vap, epsh, ep, bipsh, bip, zupsh, zup, zeepsh, zeep; vashp, vash, eshp, esh, bishp, bish, zushp, zush, zeeshp, zeesh.

(c) Other fillers: In each group (separated by a semicolon), the actual stimuli that are heard by the participant are not in parentheses. The words in parentheses are not actually presented aurally but are presented as choices in response to the stimulus that is not in parentheses in that particular group. The second choice on the screen is always the actual stimulus, which is not in parentheses. Where “2x” is noted, this indicates that the choice in parentheses was given twice, with the order in which that choice was presented on the screen relative to the correct choice differing for each of the two presentations.

beelv, (bilv), (beelf), (beeliv), (deelv); pritr, (pritr), (britv), (pritr), (pretr), (pritr); dlo, (glo), (dilo), (gilo); tloni, (tiloni), (tlomi), (dloni); est, (esk), (ezt), (esd), (ast); jisp, (chisp), (jisk), (jizb), (jesp); vorner, (forner), (vormer), (borner), tsirin, (zirin), (tisirin), (tsirin), (sirin), bugp, (bukp), (pugp), (pukp), (bugp), (bogp), (bugip), razm, (razim), (rasm), (rizm), (razim), (razn); razim, (razm), (rasim), (rizim), (razin); pab, (bab), (bap), (peb); sbo, (sibo), (zbo), (sbow), (spo); mlo, (nlo), (milo), (melo); blo, (plo), (bilo), (belo); tlon, ((tilon), (tlom), (itlon), (tlun); klon, (kilon), (klom), (iklon), (klun), kloni, (kiloni), (klomi), (itiloni), (itlomi), (utloni), stib, (stip), (step), (sdib), (stib), (step), (sdip), spib, (spip) 2x, (sbib); skib, (skip), (sikib), (skeb), zkip, (skip) 2x, (zikip), (zkik), kfite, (kvite), (kfide), (kifite), (gvite); shpik, (chpik), (shipik), (shpip), (shbik); shbik, (chbik), (shibik), (shbig), (shpik); hlone, (hilone), (hlome), (hilome); nlone, (nilon), (nlom),

(mlon); chren, (chiren), (chen), (chrem), srene, (sirene), (srem), (sen); jren, jiren, jrem, jen; sfene, (feen), (sifene), (svene); chsaf, (saf), (chaf), (chisaf), (chsav); krosk, (krost), (krozk), (kirosk), (krusk); nimsk, (mimsk), (minsk), (nimisk), (mimisk), brizm, (brizim), (brisim), (brizn), (birizm), zbrug, (zibrug), (sbrug); streeg, (sitreeg), (stireeg), (streek), (sdreeg), sgrif, (sigrif) 2x, (sgriv), (skrif); sdrig, (sidrig), (sdirig), (strig), (sdrik); dasb, (dasp), (dasib), (dazb), (dasd); digal, (dikal), (tigal); vilm, (film), (viln), (vim), (vin); zlum, (zilum), (slum), (zlun); nlim, (mlim), (nilim), (milim), fleegim, (fileegim), (fleeshim), (fleegin); vgow, (fgow), (vigow), (vkow); jrow, (jirow) 2x; tlomy, (tilomy), (tlony); vreemy, (vireemy), (vreeny), (vrimmy), dleemy, (gleemy) 2x, (gleeny) 2x, vetter (with flap), (vedder) 2x, (fetter); jeeter (with flap), (jeeder) 2x, (cheeter), ditter (with flap), (didder) 2x; nitter (with flap), (nidder) 2x, pitter (with flap), (pidder) 2x; wap, (wab) 2x, (wak) 2x; wat, (wad) 2x, (wak) 2x; jat, (jad) 2x, (jak) 2x, jlat, (jilat) 2x, (jlad), brap, (brab) 2x, (birap); grap, (girap) 2x, (grab), hlat, (hilat) 2x, (hlak); ksap, (kisap) 2x, (ksab) 2x; kack, (kag) 2x, (gack) 2x; gack, (gag) 2x, (kack) 2x; yat, (iyat), (yak) 2x, (yap); vap, (fap) 2x, (vab), (vab); yat, (iyat), (yak) 2x, (yap); vap, (fap) 2x, (vab) 2x; ijat, (jat), (ijak); usup, (uzup) 2x, (usub) 2x; ivap, (ifap) 2x; kack, (kag) 2x, (gack) 2x; gack, (gag) 2x, (kack) 2x

Experiment 2

Target stimuli: Attested clusters

gisty, bisty, kesty, wusty, shasty, gispy, bispy, kespy, wuspy, shaspy, kifty, difty, keyty, vafty, zafty, kifsy, vifsy, gefsy, vufsy, zafsy, gindy, vindy, gendy, vundy, zandy, ginty, vinty, genty, vuntty, zantty, zildy, vildy, geldy, vuldy, zaldy, zilty, vilty, klelty, vulty, zalty, zilby, vilby, gelby, vulby, zalby, zilmy, shilmy, kelmy, wulmy, zalmy, zilvy, vilvy, gelvy, vulvy, zalvy gickty, zickty, geckty, vuckty, chackty, gicksy, shicksy, gecksy, vucksy, chacksy

Target stimuli: Unattested clusters

gickpy, zickpy, geckpy, vuckpy, chackpy, gickfy, zickfy, geckfy, vuckfy, chackfy, kizby, dizby, kezby, vuzby, shazby, kizdy, dizdy, kezdy, vuzdy, shazdy, zimgy, fimgy, gemgy, zamgy vumgy, zimky, fimky, gemky, vumky, zamky, zilgy, vilgy, gelgy, vulgy, zalgy, zimjy, fimjy, gemjy, vumjy, zamjy, gipshy, bipshy, kepshy, vupshy, shapshy

Fillers:

sleepy, talky, kessy, wussy, shappy, kippy, dippy, keffy, wuffy, fishy, gitty, bitty, ketty, wutty, shatty, kitty, sappy, zoosy, shluffy, shaffy, kibby, dibby, kezy, vubby, shabby, kivy, bibby, gevy, fuchy, shavy, kiggy, picky, lowny, clowny, acty, slappy, poundy, walky, wuggy, rippy, washy, rappy, runny, fanny, zany, dilly, matty, klelly, vully, zatty, gicky, vicky, getty, vucky, chacky, flingy, shansy, geddy, vunny, zaddy, gatty, catty, gelly, pelty, singy, pushy, jokey, micky, drinky, roasty, lispy, sicky, husky, noopy, filmy, zimmy, fimmy, stingy, vummy, zammy, ziffy, viffy, geffy, vuffy, zaffy, ziggy, panty, putty, fifty, bulby, mippy, traipsy, boldy, windy, kilny, baby, bobby, vushy, patty, cryy, jippy, fuppy, zushy, goopy, sooky, chunny, zayly, jassy, juppy, shooky, spooky, dilly, divy, stupsy, barry, zijy, mijy, gejjy, vujy, zajy

Bibliography

- Akers, Glenn Alan (1981). *Phonological Variation in the Jamaican Continuum*. Ann Arbor: Karoma.
- Albright, Adam (2006) Why eatees are not E.T's: Blocking of aspiration by output-output constraints. Paper presented at the 14th Manchester Phonology Meeting (Manchester, UK, 25 May 2006).
- Archambault, Danièle and Dumouchel, Pierre (1993). *Lecture de spectrogrammes : français québécois*. Montréal: Librairie de l'Université de Montréal.
- Baek, Eung-Jin (1991). Unreleasing in Korean: a phonetic explanation. *Harvard Studies in Korean Linguistics* 4. 29–35.
- Barnes, Jonathan (2002) Positional Neutralization: A Phonologization Approach to Typological Patterns. Doctoral dissertation, University of California, Berkeley.
- Bayley, Robert (1995). Consonant cluster reduction in Tejano English. *Language Variation and Change*, 6:303-326.
- Blevins, Juliette (2004). *Evolutionary Phonology: The Emergence of Sound Patterns*. Cambridge: Cambridge University Press.
- Simo-Bobda, A. (1994). *Aspects of Cameroon English Phonology*. Berne: Peter Lang.
- Boersma, Paul (1998) *Functional Phonology: Formalizing the Interaction Between Articulatory and Perceptual Drives*. The Hague: Holland Academic Graphics.
- Cho, T. (1999) Intra-dialectal variation in Korean consonant cluster simplification: A stochastic approach. *CLS* 35,43-58.
- Chomsky, N. & M. Halle (1968). *The Sound Pattern of English*. New York: Harper and Row.
- Chung, Kook (1986). The role of [±release] and CV-constraints in phonology: evidence from Korean. In *The Linguistic Society of Korea (ed.) Linguistics in the morning calm* 2. Seoul: Hanshin. 317–336.
- Côté, Marie-Hélène (2004) Syntagmatic distinctness in consonant deletion. *Phonology* 21: 1-41
- Côté, Marie-Hélène (2000). *Consonant cluster phonotactics: A Perceptual Approach*. Ph.D. dissertation, MIT. Distributed as *MIT Working Papers in Linguistics and Philosophy*.
- Katzir Cozier, Franz (2008). Encoding of Perceived Contrast Between CC-Clusters and their Simplified Counterparts in Coda Cluster Simplification. Papers presented at WCCFL 27, UCLA 2008, and the Consonant Clusters and Structural Complexity Workshop, Munich 2008.
- Dorman, M.F., M. Studdert-Kennedy, and L.J. Raphael (1977). Stop-consonant recognition: Release bursts and formant transitions as functionally equivalent, context-dependent cues. *Perception and Psychophysics* 22: 109-22.
- Flemming, Edward (2004). Contrast and perceptual distinctiveness. Bruce Hayes, Robert Kirchner, and Donca Steriade (eds.) *The Phonetic Bases of Markedness*. Cambridge University Press, Cambridge. 232-276.
- Flemming, Edward (2006). The role of distinctiveness constraints in phonology. Massachusetts Institute of Technology, ms.

- Goeman, A.C.M. (1999) *T-deletie in Nederlandse dialecten. Kwantitatieve analyse van structurele, ruimtelijke en temporele variatie*. Holland Academic Graphics, Den Haag
- Green, Lisa (1992). Consonant cluster reduction and suffix interface in African american English. University of massachusetts, Amherst, ms.
- Hayes, Bruce, Bruce Tesar and Kie Zuraw (2003). *OTSoft 2.1*. Los Angeles and New Brunswick, NJ: UCLA and Rutgers University. [Software package available at <http://www.linguistics.ucla.edu/people/hayes/otsoft/>].
- Hinskens, Frans (1992). *Dialect Leveling in Limburg. Structural and Sociolinguistic Aspects*. Ph.D. dissertation, Katholieke Universiteit te Nijmegen.
- Hinskens, Frans (1996). *Dialect Leveling in Limburg: Structural and Sociolinguistic Aspects*. Tübingen: Niemeyer.
- Kaisse, Ellen M. and Patricia A. Shaw. 1985. On the theory of Lexical Phonology. *Phonology Yearbook*, 2:1-30.
- Huh, Woong (1965). *Kwukeumwunhak*. [Korean phonology.] Seoul: Cengumsa.
- Hura, S.L., Lindblom, B., & Diehl, R.L. (1992). On the role of perception in shaping phonological assimilation rules. *Language and Speech*, 35, 59-72.
- Hyman, Larry (2001) The limits of phonetic determination in phonology: *NC revisited. In *The Role of Speech Perception in Phonology*, eds. E. Hume and K. Johnson, 141-185. San Diego: Academic Press.
- Jun, Sun-Ah (1993). The phonetics and phonology of Korean prosody. PhD dissertation, Ohio State University.
- Kang, Yoonjung (2003) Perceptual similarity in loanword adaptation: English postvocalic word-final stops in Korean Phonology 20 219–273.
- Kaplan, Abby (2006) English Postnasal stop deletion: *NT vs. *ND. Paper presented at Phorum, UC Berkeley (November 20, 2006).
- Kewley-Port, D., D.B. Pisoni, and M. Studdert-Kennedy (1983). Perception of static and dynamic acoustic cues to place of articulation in initial stop consonants. *JASA* 73: 1779-93.
- Kim, Chin-Wu (1971). Two phonological notes: A-sharp and B-flat. In Michael Brame (ed.) *Contributions to generative phonology*. Austin: University of Texas Press. 155–170.
- Kim, Hyunsoon (1998). A phonetic characterization of release and nonrelease: the case of Korean and English. *Language Research* 34. 347–368.
- Kim-Renaud, Young-Key (1974). Korean consonantal phonology. PhD dissertation, University of Hawaii.
- Kohler, K. (1990) Segmental reduction in connected speech: Phonological facts and phonetic explanations. In Hardcastle, W.J.; Marchal, A. (eds.): *Speech Production and Speech Modeling*, pp. 69-92. Kluwer Academic Publishers, Dordrecht.
- Lee, Sechang (1994). Neutralization and tensification in Korean. In Noriko Akatsuka (ed.) *Japanese/Korean linguistics 4*. Stanford: CSLI. 511–520.
- Lisker, Leigh (1999). Perceiving final voiceless stops without release: Effects of preceding monothongs versus nonmonothongs. *Phonetica* 56; 44-45.
- Macmillan, N.A., & Creelman, C.D. (2006). *Detection Theory: a user's guide* (2nd ed.). Mahwah, NJ: elrbaum.
- Martin, Samuel E. (1951) Korean Phonemics, *Language* 27, 519-533.
- Mascaró, Joan (1978). Catalan Phonology and the Phonological Cycle. Doctoral dissertation, MIT.

- McCawley, James D. (1967). Le rôle d'un système de traits phonologiques dans un théorie du langage. *Langages* 8. 112–123. Reprinted as 'The role of a phonological feature system in a theory of language' in Valerie Makkai (ed.) (1972). *Phonological theory*. New York: Holt, Rinehart & Winston. 522–528.
- Meade, Rocky R. (2001). *Acquisition of Jamaican phonology*. Dordrecht, Foris, Thesis UvA. LOT International Series, 41.
- Miller, George A. and Nicely, Patricia. (1955). An Analysis of Perceptual Confusions among some English Consonants, *Journal of the Acoustical Society of America* 27:2.
- Munson, Benjamin (2001). A method for studying variability in fricatives using dynamic measures of spectral mean. *Journal of the Acoustical Society of America* 110, 2: 1203–1206
- Naeser, M.A. (1970). Influence of initial and final consonants on vowel duration in CVC syllables. *Technical Report No. 130*, Wisconsin Research and Development Center for Cognitive learning, The University of Wisconsin.
- Ohala, John (1981) The listener as a source of sound change. In Proceedings of Chicago Linguistic Society: Papers from the Parasession on Language and Behavior, 178-203. Chicago: CLS.
- Ohala, J.J. (1990). The phonetics and phonology of aspects of assimilation. In *Papers in laboratory phonology I: between the grammar and the physics of speech* (J. Kingston and M. Beckman, eds.), 258-275. Cambridge: Cambridge University Press.
- Patrick, Peter L. (1991). "Creoles at the intersection of variable processes: -t, -d deletion and past-marking in the Jamaican mesolect." *Language Variation & Change* 3(2): 171-189.
- Rhee, Seok-Chae (1998). Aspects of release and nonrelease in phonology. PhD dissertation, University of Illinois at Urbana-Champaign.
- Schouten, M.E.H. (1982). T-deletie in de stad Utrecht: schoolkinderen en grootouders. *Forum der Letteren*, 23:282-291.
- Schouten, M.E.H. (1984). T-deletie in het zuiden van die provincie Utrecht. *Taal en Tongval*, 36:162-173.
- Sohn, Hyang-Sook (1987). Underspecification in Korean phonology. PhD dissertation, University of Illinois at Urbana-Champaign.
- Steriade, D. (1997). Phonetics in Phonology: The Case of Laryngeal Neutralization. ms. UCLA.
- Steriade, D. (2001a). Directional asymmetries in place assimilation: a perceptual account. In E. Hume and K. Johnson (eds.) *The Role of Speech Perception in Phonology*. New York: Academic Press.
- Steriade, D. (2001b). The Phonology of Perceptibility Effects: the P-map and its consequences for constraint organization. UCLA ms.
- Wheeler, M. 1979. *Catalan Phonology*. Oxford: Basil Blackwell.
- Wright, Richard (2004). 'A review of perceptual cues and cue robustness'. In Bruce Hayes, Robert Kirchner, and Donca Steriade, eds., *Phonetically-Based Phonology*, Cambridge University Press.